Genome-wide analysis of the transcriptional network controlled by PGC-1α and ERRα in skeletal muscle from microarray and next generation sequencing data

Inauguraldissertation

zur
Erlangung der Würde eines Doktors der Philosophie
vorgelegt der
Philosophisch-Naturwissenschaftlichen Fakultät
der Universität Basel

von

Silvia Salatino
aus Italien

Basel, 2014

Genehmigt von der Philosophisch-Naturwissenschaftlichen Fakultät

auf Antrag von

Prof. Dr. Christoph Handschin und Prof. Dr. Erik van Nimwegen

Basel, den 24.06.2014

Prof. Dr. Jörg Schibler, Dekan

Table of contents

Summary

Gene regulation in higher eukaryotes is a complex and dynamic process involving the coordinated action of several transcription factors and coregulators. The assembly of multi-protein complexes at promoter or enhancer elements modulates the activity of whole biological pathways and contributes to the physiological plasticity of cells and tissues. However, given their complexity and the high number of players involved, these regulatory processes are still poorly understood. In this context, next generation sequencing methods like the chromatin immunoprecipitation sequencing technique (ChIP-Seq) provide a novel approach for mapping the interactions between protein complexes and DNA elements on a genome-wide scale.

In this project, we perform a comprehensive coregulator-tailored ChIP-Seq data analysis to study the global recruitment of the peroxisome proliferator-activated receptor γ coactivator 1 α (PGC-1α) and of the nuclear estrogen-related receptor α (ERRα) in skeletal muscle. Moreover, by integrating this information with microarray expression data, we infer the direct and indirect effects of our proteins of interest on their downstream target genes and, thus, on the regulated biological pathways. In addition, by combining several computational techniques, including binding site prediction and principal component analysis, we identify the activator protein 1 (AP-1) and the specificity protein 1 (SP1) as novel transcriptional partners in the PGC-1α- and ERRα-mediated regulation of energy metabolism in skeletal muscle. Our study provides a new approach for the genome-wide analysis of coregulators and sheds light on the transcriptional network controlling skeletal muscle plasticity.

The content of this thesis is organized as follows. In chapter 1 we depict the state-of-the-art regarding nuclear receptors (and in particular ERRα), coregulators (with a focus on PGC-1α) and next generation sequencing technologies; moreover, we present the aims that motivated the projects described in this thesis. Chapter 2 illustrates our ChIP-Seq data analysis procedure, starting from the raw reads till the peak calling step. As every maturing technique, ChIP-Seq is accompanied by a number of issues that still need to be addressed; for this reason, we focus particularly on the artifacts we found and on our computational approaches to solve them. In chapter 3 we describe the downstream analysis steps of ChIP-Seq studies, including peak annotation, motif search and principal component analysis of peak binding site predictions. The results of the computational techniques described in the previous chapters to dissect PGC-1α regulatory network in skeletal muscle are illustrated in chapter 4, whereas in chapter 5 we focus on the interplay between PGC-1α and ERRα and in particular we examine to which extent ERRα

is required for PGC-1α-mediated effects on skeletal muscle cells. Finally, the discussion will review the main findings of the present thesis and provide an outlook on the possible future developments of this study.

Abbreviations

ACADM	medium-chain specific acyl-CoA dehydrogenase
ACTR	activator of thyroid and retinoic acid receptor
AD	activation domain
AF-1	activation function domain 1
AF-2	activation function domain 2
AMPK	AMP-activated protein kinase
AP-1	activator protein 1
AR	androgen receptor
AV	adenovirus
BAF60a	BRG1-associated factor 60a
BAT	brown adipose tissue
bp	basepair
cAMP	cyclic adenosine monophosphate (AMP)
CEBPB	CCAAT/enhancer-binding protein β
ChIP-Seq	chromatin immunoprecipitation followed by deep sequencing
CNV	copy-number variant
COX4	cytochrome c oxidase subunit 4
CpG	cytosine-phosphate-guanine dinucleotide
CREB	cyclic AMP response element-binding protein
CRM	cis-regulatory module
Ctrl	control
Cycs	cytochrome C, somatic
DBD	DNA binding domain
DES	diethylstilbestrol
DNA	deoxyribonucleic acid
ERα	estrogen receptor α
ERRα	estrogen-related receptor α
ERRβ	estrogen-related receptor β
ERRγ	estrogen-related receptor γ
FC	fold change
FDR	false discovery rate
FOXA1	forkhead box protein A1
FXR	farnesoid X receptor
GABPA	GA-binding protein α chain
GCN5	general control of amino acid synthesis protein 5
GEO	Gene Expression Omnibus
GFP	green fluorescent protein
GO	Gene Ontology
GR	glucocorticoid receptor
GRIP1	glucocorticoid receptor interacting protein 1
HAT	acetyltransferase
HSF1	heat shock factor 1
IP	immunoprecipitation
ISWI	imitation SWI protein
kb	kilobase
KD	knockdown
KEGG	Kyoto Encyclopedia of Genes and Genomes
LBD	ligand binding domain
LRH-1	liver receptor homolog-1
LXR	liver X receptor
MAPK	mitogen-activated protein kinase
MARA	Motif Activity Response Analysis

MKO	muscle-specific knockout
mTOR	mammalian target of rapamycin
NCoR1	nuclear receptor-corepressor 1
NF-κB	nuclear factor-κB
NGS	next generation sequencing
NR	nuclear receptor
NRF-1	nuclear respiratory factor 1
NRF-2	nuclear respiratory factor 2
NRRE	nuclear receptor response element
OXPHOS	oxidative phosphorylation
OV	over-expression
PAR-CLIP	photoactivatable-ribonucleoside-enhanced crosslinking immunoprecipitation
PCA	principal component analysis
PCR	polymerase chain reaction
PDK4	pyruvate dehydrogenase kinase 4
PGC-1α	peroxisome proliferator-activated receptor γ coactivator 1 α
PGC-1β	peroxisome proliferator-activated receptor γ coactivator 1 β
PPAR	peroxisome proliferator-activated receptor
PRC	PGC-1-related coactivator
PROX1	prospero-related homeobox 1
PTM	post-translational modification
PWM	position weight matrix
qRT-PCR	quantitative reverse transcription polymerase chain reaction
RCoR	REST corepressor protein
RIP140	receptor-interacting protein 140
RNA	ribonucleic acid
RNA-Seq	RNA sequencing
ROS	reactive oxygen species
RSC	remodeling the structure of chromatin complex
RXR	retinoid X receptor
SEM	standard error of the mean
SF1	steroidogenic factor 1
SHP	small heterodimer partner
shRNA	small hairpin RNA
siRNA	small interfering RNA
SIRT1	sirtuin 1
SMRT	nuclear receptor-corepressor 2 (NCoR2)
SNP	single-nucleotide polymorphism
SP1	specificity protein 1
SRC-1	steroid receptor coactivator 1
SVD	singular value decomposition
SWI/SNF	SWitch/Sucrose NonFermentable
TAM	tamoxifen
TES	transcription end site
TF	transcription factor
TFAM	mitochondrial transcription factor A (also named mtTFA)
TFBS	transcription factor binding site
TG	transgenic
TSS	transcription start site
UFE	unknown functional element
VEGF	vascular endothelial growth factor
WCE	whole cell extract
WM	weight matrix

1. Introduction

1.1. Nuclear receptors

Nuclear receptors (NRs) are a class of transcription factors that modulate several major physiological processes, ranging from development to reproduction, inflammation, and metabolism [1]. As for all transcription factors, this regulatory capacity of NRs relies on their ability to recognize specific DNA sequences at the promoter of their target genes and to affect the rate of RNA polymerase II-directed transcription by interacting with other components of the general transcriptional machinery [2]. The next sections will provide a detailed description of NR different types and functions, as well as an in-depth portrait of the ERR family of nuclear receptors and in particular of ERRα.

1.1.1. Classification, structure and function

The nuclear receptors superfamily is encoded by a set of 49 genes in mouse (48 in human and 47 in rat) [3] and consists of two distinct groups of transcription factors: the first includes the "ligand-activated" NRs, which transduce thyroid hormones, steroid hormones, retinoids and vitamin D signals into physiological responses; the second, namely the "orphan receptors", is a group of factors for which no known endogenous ligands have been identified so far [4]. Because of their specific response to small molecules, the ligand-induced family of NRs has been more extensively studied to develop drugs for the treatment of several pathological states, including obesity, diabetes, cancer and cardiovascular or metabolic diseases; the growing interest in these therapeutic targets has led to the creation of numerous databases and web resources containing the set of all transcription factor binding sites on a genome-wide scale, also called "cistrome" (for some examples, see [5, 6]).

As illustrated in Figure 1.1, the domain structure common to most of the NR superfamily members consists of four functional units: an amino-terminal activation domain (named AF1), a central DNA binding domain (or DBD) and a carboxy-terminal ligand binding domain (also known as LBD), which hosts a second activation function domain (AF2). The most conserved region is the DBD, which consist of two zinc finger domains to recognize specific NR-responsive elements (NRREs) in the regulatory regions of their target genes [7]. The LBD domain, whose function is to provide a hydrophobic pocket for binding recognition, is also a target for post-translational

modifications, which alter the NR stability and binding affinity, resulting in a modulated target gene expression.

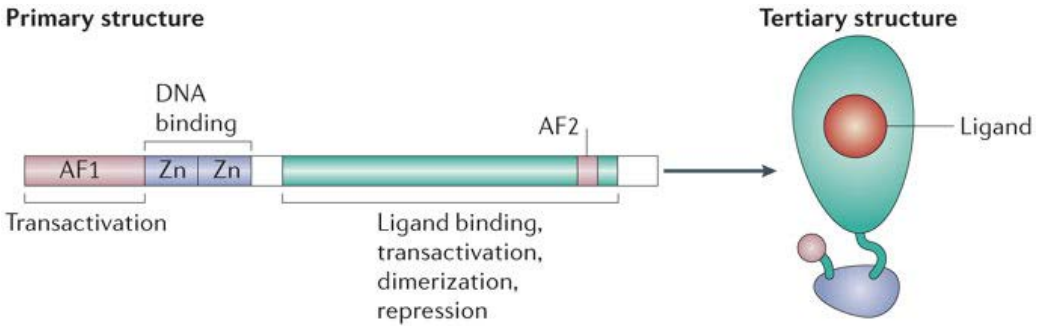


Figure 1.1 Nuclear receptors structural domains. *From [8].*

Nuclear receptors can be classified according to several criteria; for example, on the basis of the ligand-binding properties, NRs can be subdivided in hormone receptors, metabolic receptors and orphan receptors (Fig.1.2). NRs belonging to the first group – such as the androgen receptor (AR) or the glucocorticoid receptor (GR) – generally bind to the DNA as homodimers and are usually located in the cytosol; upon binding to their ligands, they translocate to the nucleus and recognize their cognate NRRE sequences on the DNA. Metabolic receptors – like for example the farnesoid X receptor (FXR), the liver X receptor (LXR) and the peroxisome proliferator-activated receptors (PPARs) – are instead located in the nucleus and often bound to the DNA in heterodimers with the retinoid X receptor (RXR). The third group of receptors – which include the steroidogenic factor 1 (SF1) and the liver receptor homolog-1 (LRH-1) – are generally regulating transcription through changes in their expression or post-translational modifications and can work either as heterodimers or as monomers [7].

Chromatin accessibility to DNase I and histone modifications contribute to nuclear receptors binding to the DNA both before and after the interaction with their cognate ligands: John and colleagues have shown that the glucocorticoid receptor often binds to pre-existing regions of open chromatin [9], while Hurtado et al. reported that a significant fraction of estrogen receptor alpha sites were accessible before estrogen induction [10]. On the other hand, chromatin remodeling continues also post-induction, as it has been shown that the binding of certain nuclear receptors, as well as some nuclear factors – like the forkhead box protein A1 (FOXA1) – can act as “pioneer factors” to facilitate the binding for a second NR through recruitment of nucleosome remodeling complexes (such as the SWitch/Sucrose NonFermentable complex SWI/SNF), in a process called “assisted loading” [11, 12].

Nuclear receptors can regulate transcription, both positively and negatively, by multiple mechanisms. In the absence of a ligand, NRs are often part of complexes including corepressor proteins such as the nuclear receptor co-repressor 1 (NCoR1) or the nuclear receptor co-repressor 2 (also called SMRT). However, this ligand-independent repression is not the only way to diminish target genes expression: several NRs (e.g. GR or LXRs) can also repress transcription in a ligand-dependent manner (named "transrepression") by antagonizing and inhibiting the activity of other signal-dependent transcription factors; one example is the nuclear factor-κB (NF-κB), which plays a crucial role as effector of pro-inflammatory signaling pathways. Differently from other transcription factors, several NRs are able to activate transcription only upon the binding of their cognate ligand, which: (i) triggers the dissociation of the corepressor, (ii) induces a conformational change in the α-helical region of the C terminal domain AF2 and (iii) stabilizes the binding of the receptor with the DNA. This ligand-dependent transactivation has the role of recruiting the coactivator proteins and, through these, of modifying the local chromatin structure to facilitate the assembly of the transcriptional complex [13].

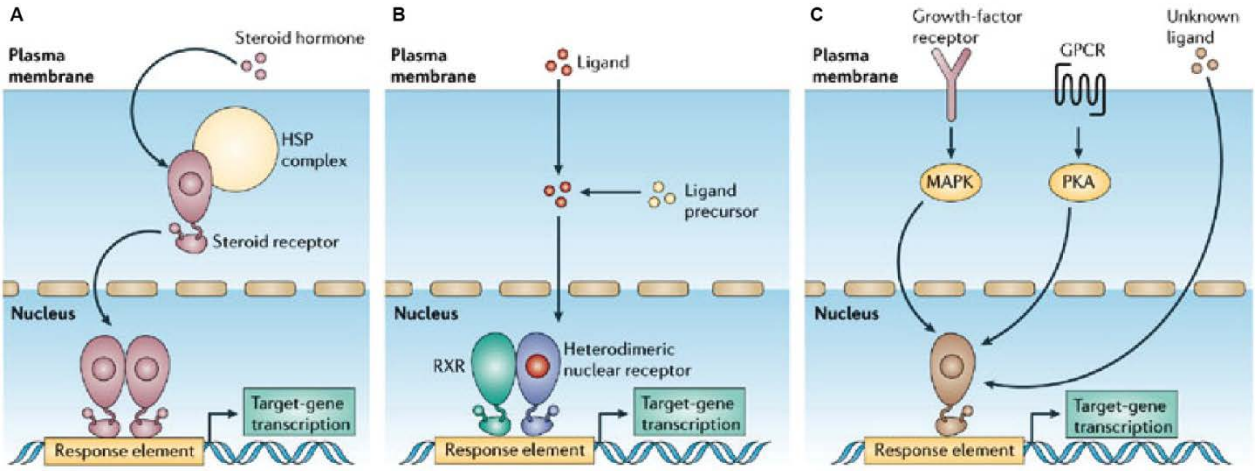


Figure 1.2 Nuclear receptors can be classified in three groups on the basis of their ligand-binding properties: (**A**) hormone receptors, (**B**) metabolic receptors and (**C**) orphan receptors. *From [8].*

1.1.2. The ERR family of nuclear receptors

One interesting group of nuclear receptors is represented by the estrogen-related receptor (ERR) subfamily. All the three members of this family, ERRα, ERRβ and ERRγ belong to the orphan receptor category, as no endogenous ligands have been identified to date. Based on its high sequence homology with the estrogen receptor alpha (ERα), ERRα was the first of the three NRs, together with ERRβ, to be identified [14]. However, differently from ERα, neither ERRα nor the other two members of the family were shown to bind estrogen [14], as confirmed

by crystallographic and functional studies indicating that the ERRs are constitutively active without a natural ligand [15, 16]. On the other hand, several synthetic ligands have been found to modulate the activity of the ERRs. One of these is the synthetic nonsteroidal estrogen diethylstilbestrol (DES), widely used to prevent spontaneous abortion in women until it was withdrew by the United States Food and Drug Administration in the 70s, as it was shown to cause a rare vaginal tumor. Other synthetic ligands are tamoxifen (TAM) and its derivative 4-OHT, which act as estrogen antagonists in mammary glands and are clinically used for breast cancer treatment. Although the in vitro effects of DES and TAM were documented for ERRβ and ERRγ, no endogenous ERRα-interacting molecule has been identified to date [17], while the synthetic compound XCT-790 has been shown to act as a selective inverse agonist of ERRα [18].

The DNA sequence recognized by all the three family members is the extended half site TNAAGGTCA, which is present either single or repeated and to which the ERRs can bind as monomers or dimers, respectively; in the latter case, although several groups have suggested that the orientation and spacing between the two half sites might determine the specificity of recognition, an in-depth study on ERRγ reported instead no clear relationships between the spacing and the binding affinity of the nuclear receptor [19]. Furthermore, Zhang and others have demonstrated that ERRs can bind and compete with ERα to the classical estrogen response element sites (EREs) [20], which is a 13 bp-long palindromic sequence composed by two inverted repeats separated by a 3 bp-long central spacer [21].

Concerning the biological functions of ERRs, a number of studies have established a central role of these nuclear receptors in the regulation of energy metabolism and mitochondrial biogenesis [22], as confirmed by their expression patterns: in fact, although they are ubiquitously expressed, their level is much higher in tissues with high energy demands, like brain, heart, kidney, retina, brown adipose tissue and skeletal muscle [23]. Albeit they have overlapping functions, the ERRs carry also distinct biological roles. For example, ERRβ plays crucial roles in placental development [24], inner ear maturation [25], photoreceptor survival [26] and type I muscle fiber formation together with ERRγ [27]. On the other hand, ERRγ was shown by a loss-of-function study to be specifically important for the maintenance of cardiac, gastric, and renal potassium homeostasis [28]. Moreover, ERRγ muscle-specific transgenic mice showed increased expression of genes promoting fat metabolism, TCA cycle and vascularization [29].

Similarly to ERRγ, ERRα also plays a relevant role in the regulation of energy metabolism and in particular of fatty acids metabolism, as it induces the pyruvate dehydrogenase kinase 4 (PDK4),

a negative regulator of glucose oxidation [30]. Although no phenotypic change in skeletal muscle (apart from reduced fat mass) has been observed in whole-body ERRα knockout animals [31], the over-expression of ERRα in C2C12 myoblasts accelerated their differentiation to myotubes via modulation of the mitogen-activated protein kinase (MAPK) signaling pathway, highlighting the importance of ERRα in promoting muscle growth [32]. Several studies have shown that the peroxisome proliferator activated receptor γ coactivator 1α (PGC-1α) is a strong inducer and coactivator of ERRα, which in turn promotes its own transcription in a positive feedback loop and regulates mitochondrial function [33-35].

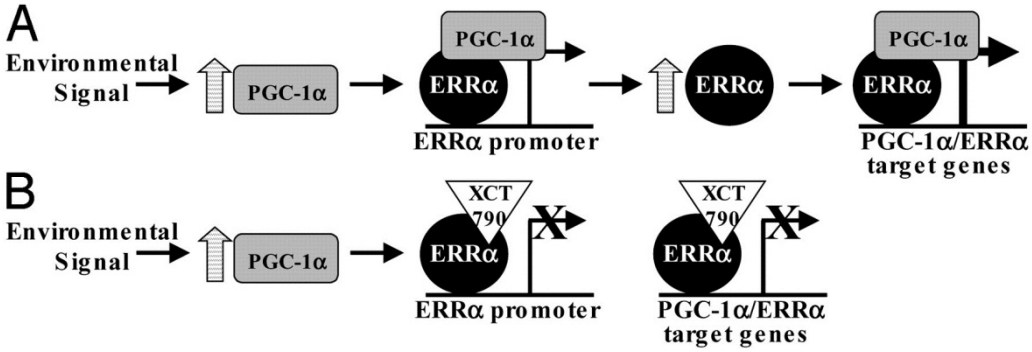


Figure 1.3 Schematic model for regulation of ERRα and PGC-1α target genes. *Adapted from [36]*.

1.2. Coregulators of transcription

The efficacy of nuclear receptor-mediated transcription strongly depends on coregulators, a group of proteins that associate with NRs and alter the state of the chromatin to either enhance ("*coactivators"*) or inhibit ("*corepressors"*) NRs ability of regulating the expression of hormone-responsive genes [37]. The abundance of these coregulators – which include also the receptor-interacting protein 140 (RIP140) and the nuclear receptor co-repressor 1 (NCoR1) – is regulated in response to a variety of physiological stimuli, (such as exercise, cold and fasting) and in pathological conditions [38].

1.2.1. Types and molecular mechanism of action

Apart from their final effect as activators or repressors of transcription, coregulators can be distinguished on the basis of their mechanism of action in two main groups: the ATP-dependent remodeling factors and the histone acetyltransferases (HATs).

Coregulators belonging to the first group use the energy produced from the hydrolysis of ATP to introduce a superhelical torsion into the chromatin, thus actively altering nucleosome positions (by sliding or ejecting them) and rendering the DNA either more or less accessible to transcription factors [39]. The yeast SWI/SNF complex was the first remodeling factor to be identified, followed by the imitation SWI (ISWI) proteins and the RSC (remodeling the structure of chromatin) complex, whose corresponding human orthologs are the SMARC proteins.

Many coactivator proteins either possess intrinsic histone acetyltransferase enzymatic activities or recruit other proteins with this activity to regulatory regions. At physiological pH, the electrostatic interactions between the deprotonated phosphates of the DNA backbone (hence, positively charged) and the protonated lysine residues of the histones (negatively charged) are responsible for the binding of the DNA to the histones. The function of the HAT proteins is to acetylate the side chains of the histone lysines, thus weakening the tight bond that keeps the DNA wrapped around them. As a consequence, the chromatin becomes more accessible to transcription factors and to the general transcriptional machinery. Conversely, several corepressor proteins possess histone deacetylase properties or recruit histone deacetylases (HDACs) to remove the acetyl groups from lysines and, thus, restoring the condensed and inaccessible state of the DNA. Some examples of corepressors belonging to this group are the REST corepressor proteins (RCoR 1, 2 and 3) and the nuclear receptor co-repressor proteins (NCoR1 and 2); in skeletal muscle, NCoR1 can repress ERRα under basal conditions, thereby decreasing metabolic genes transcription, while it is significantly reduced upon exercise and caloric restriction [40, 41]. On the other hand, examples of coactivator proteins include the p300 protein, which plays an essential role in regulating cell growth, and the peroxisome proliferator activated receptor γ coactivator 1 proteins (PGC-1α and β).

1.2.2. The PGC-1 family of coactivators

The peroxisome proliferator activated receptor (PPAR) γ coactivator 1 (PGC-1) protein family is composed of three members: PGC-1α, PGC-1β and the PGC-1-related coactivator (PRC). The structure features common to all the three proteins are: (i) an N-terminal activation domain (AD), (ii) one or more LXXLL or LLXXL motifs (where L stands for leucine and X for any amino acid), which are used to interact with nuclear receptors, (iii) a serine-rich domain and (iv) a C-terminal RNA recognition and splicing motif (for a detailed description, see [42]). Similarly to other coactivators, due to the lack of a DNA binding domain, the PGC-1 members do not interact directly with the DNA, but rather act as scaffolds for the recruitment of other proteins containing

enzymatic activities and of NRs [43]. With these, they form multi-protein complexes that positively regulate the expression of their target genes.

PRC can be found ubiquitously, whereas PGC-1α and PGC-1β are highly expressed in organs with high energy demands (in terms of oxygen and ATP consumption), like skeletal muscle, heart, kidney, brain and brown adipose tissue. While PRC has not been extensively studied, PGC-1β is known to play a role in muscle endurance capacity and angiogenesis [44, 45], in liver lipogenesis [46], osteoclast activation [47] and brown fat differentiation [48]. In the next sections we will focus on the modulation of PGC-1α expression and activity, as well as on the transcriptional network regulated by this coactivator in skeletal muscle.

1.2.3. Modulation of PGC-1α expression and activity

The most studied member of the PGC-1 family is undoubtedly PGC-1α. As it is a potent inducer of transcription, its regulation is tightly controlled by several mechanisms (Fig.1.4), including intracellular signaling cascades and post-translational modifications (PTMs).

Some examples of the several intracellular events that control the synthesis of PGC-1α mRNA are: (i) the cyclic AMP response element-binding protein (CREB), which binds to the PGC-1α promoter when phosphorylated upon exercise-induced calcium release [49]; (ii) the AMP-activated protein kinase (AMPK), following muscle contraction [50]; (iii) the MEF2C and MEF2D transcription factors, which trigger an auto-regulatory feed-forward loop [51]; (iv) the FOXO1 transcription factor, that inhibits PGC-1α expression in liver [52]; (v) the orphan nuclear receptor SHP, shown to reduce PGC-1α transcription in brown adipose tissue [53].

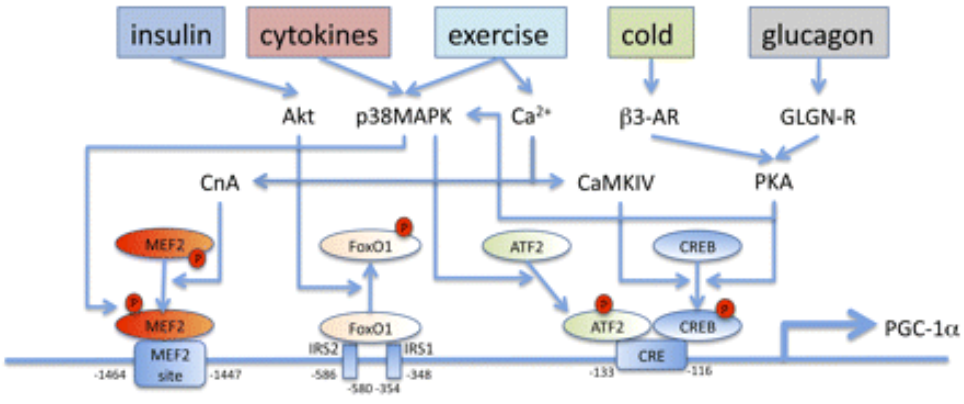


Figure 1.4 Regulation of PGC-1α transcription. The promoter of PGC-1α hosts binding sites for the transcription factors MEF2, FoxO1, ATF2 and CREB, which enhance PGC-1α transcription in response to distinct signaling pathways (insulin, cytokines, exercise, cold and blood glucagon levels). *From [54].*

Several PTMs of PGC-1α have been reported, including methylation, phosphorylation, sumoylation, acetylation and ubiquitination. For example, phosphorylation of its serine and threonine residues by AMPK and p38/MAPK are able to stabilize the relatively short half-life (2 or 3 h) of PGC-1α [50, 55]; on the contrary, poly-ubiquitination is a signal for proteosomal degradation of the protein [56]. The acetylation of PGC-1α lysine residues determines its activity: in fact, while the acetylation by the general control of amino acid synthesis protein 5 (GCN5) inactivates PGC-1α [57], the deacetylation by the protein sirtuin 1 (SIRT1) re-activates PGC-1α in response to the NAD^+/NADH ratio [58].

1.2.4. Transcriptional network regulated by PGC-1α in skeletal muscle

Originally discovered as thermogenic regulator in brown adipose tissue (BAT) upon cold exposure [59], PGC-1α has later been shown to play different roles in other tissues. For example, it is required to reduce neurodegeneration caused by reactive oxygen species (ROS) in brain [60], for protection against light-induced damages in retina [61] and for gluconeogenesis, ketogenesis, heme and bile acids biosynthesis in liver [49, 62-64].

However, the tissue in which PGC-1α shows the most remarkable effects is skeletal muscle (for an overview, see Fig.1.5). In this organ, the nodal functions of PGC-1α are the coordination of oxidative metabolism and the promotion of mitochondrial biogenesis, which are key determinants of skeletal muscle plasticity and are required to meet higher energy demands [65]. Mitochondrial biogenesis is a long-term adaptive process that results in the increase of the mitochondrial content of a tissue and in the change of mitochondrial protein-to-lipid ratio [66]. This process requires the coordinated transcription of a large number of nuclear mitochondrial genes, the synthesis of new proteins and lipids and the replication and transcription of the mitochondrial DNA, which encodes some essential components of the oxidative phosphorylation (OXPHOS) complexes [67]. In response to exercise, PGC-1α mRNA and protein levels increase, enabling this coactivator to integrate physiological signals and to enhance mitochondrial biogenesis [68, 69]. For this process, several transcription factors cooperate with and are coactivated by PGC-1α. The nuclear respiratory factor 1 (NRF-1), NRF-2 (also called GA-binding protein or GABP) and ERRα are some well-known examples; upon induction of their genes, PGC-1α stimulates the expression of mitochondrial genes including the cytochrome c oxidase subunit 4 (COX4) enzyme and the mitochondrial transcription factor A (Tfam) which, in turn, regulates transcription of the mitochondrial DNA genes [33, 68]. A direct consequence of mitochondrial biogenesis is an increase in ROS production, which is addressed by PGC-1α through the enhanced expression of the ROS detoxifying enzymes, although this has been

shown so far only for the brain [60]. Moreover, to provide an adequate energy supply to the cell, PGC-1α also promotes lipid refueling through activation of *de novo* lipogenesis and increases fatty acids oxidation by coactivating PPARα [70, 71].

PGC-1α is highly expressed in skeletal muscle and particularly enriched in the *slow-twitch* type I fibers, which are more oxidative, vascularized and rich in mitochondrial content than the *fast-twitch* type II fibers. Transgenic expression of PGC-1α showed to drive a shift in fiber type from fast to slow-twitch fibers and the transition from a resistance phenotype (i.e. with high muscle strength and mass) to a more aerobic endurance phenotype [72]. In rodents, PGC-1α was shown to be transiently induced after a single bout of low-intensity swimming exercise [73]; accordingly, chronically elevated levels of PGC-1α mRNA were observed upon endurance exercise training in humans [69]. In contrast, muscle-specific deletion of PGC-1α resulted in reduced resistance to fatigue and elevated markers of inflammation [74] , thereby providing a link between metabolic and immune signaling pathways (for an extensive review, see [42]).

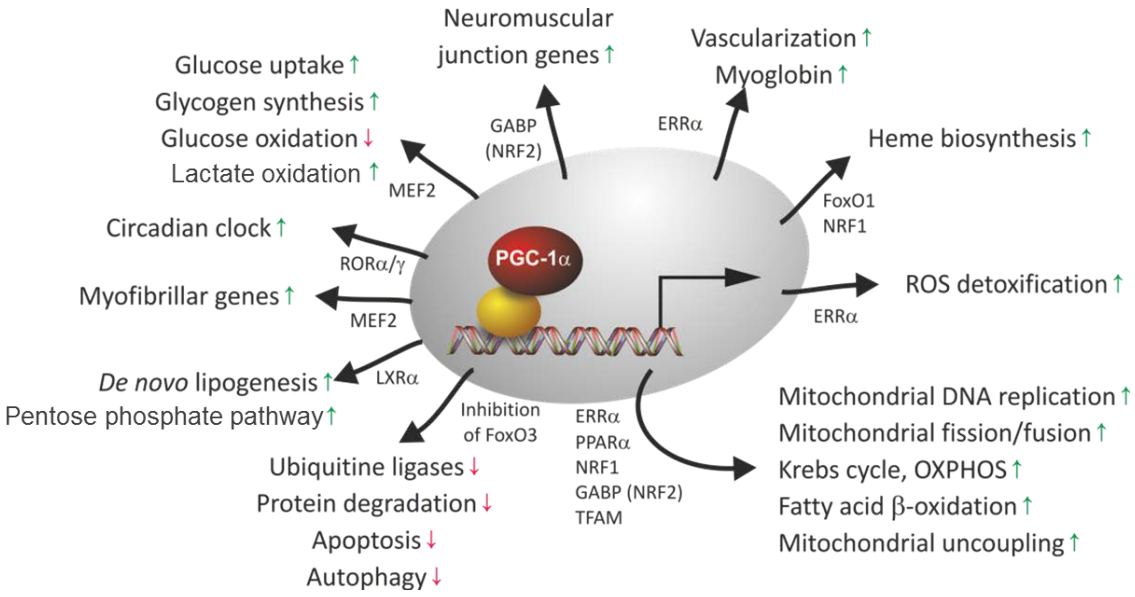


Figure 1.5 Overview of the multiple pathways in skeletal muscle regulated by PGC-1α through distinct TFs. Green and red arrows represent up- and down-regulation, respectively. *Adapted from [65].*

1.3. Next generation sequencing techniques

1.3.1. The ChIP-Seq procedure

DNA sequencing is the process used to precisely determine the exact order of nucleotides (adenine, cytosine, guanine and thymine) composing a fragment of deoxyribonucleic acid (DNA).

As this molecule encodes the genes of a living organism and the instructions to express them in time and space, the determination of its sequence has become of crucial importance in Biology for a better understanding of an organism's characteristics and diseases.

Since the "chain-termination method" was developed by Frederick Sanger in 1977, it has become the method of choice for almost three decades and, thanks to its low error rate (per-base accuracy as high as 99.999%) [75], it is still in use for small-scale experiments and to confirm *de novo* sequences. Briefly, the method consists into four separate sequencing reactions, each of which containing a chemically-altered deoxynucleotide (dATP, dCTP, dGTP, dTTP). Every time the enzyme DNA polymerase incorporates at random one of these molecules into the growing DNA chain, the copying process is stopped. The resulting DNA fragments are loaded on a gel electrophoresis chamber, where the original sequence can be deduced by reading from the smallest to the largest piece of DNA.

Over the past ten years, the "next generation sequencing" (NGS) has progressively replaced Sanger sequencing and has become the gold-standard to obtain millions of DNA sequenced fragments in a much more time- and cost-effective manner. In fact, although every NGS platform employs a different technology, the key difference with traditional Sanger sequencing is that these new methods are able to massively parallelize the sequencing process and to perform, for example, the "high-throughput" sequencing of an entire human genome in less than one day [76].

A multitude of NGS applications have arisen in the last decade. For example, NGS techniques can be used to detect large structural variations as copy-number variants (CNVs) as well as single-nucleotide polymorphisms (SNPs) and other chromosomal rearrangements [77]. Likewise, NGS methods are intensively used for whole transcriptome profiling analysis (RNA-Seq) to take a snapshot of the RNA presence and quantity at a given moment in time; mRNA studies in particular have begun to replace the use of microarrays to determine expression profiles. Other methods, like the photoactivatable-ribonucleoside-enhanced crosslinking immunoprecipitation (PAR-CLIP), can be used to analyze protein interactions with RNA and are particularly useful to detect miRNA target sites [78]. Finally, ChIP-sequencing (or ChIP-Seq) is a widespread procedure to analyze protein interactions with DNA.

By combining chromatin immunoprecipitation (ChIP) with massively parallel sequencing, ChIP-Seq can be used to precisely map binding sites for a protein of interest genome-wide. Since it was developed in 2007 [79, 80], this method has suddenly replaced the ChIP-on-chip technique

(i.e. ChIP on a microarray) thanks to the higher resolution and genome coverage. The sensitivity of this technology depends on several factors, including the sequencing platform employed and the sequencing depth (or coverage), which corresponds to the number of reads representing a given basepair in the reconstructed sequence. The method, which can be applied both to transcription factors and to epigenetic histone modifications, consists of two steps (Fig.1.6).

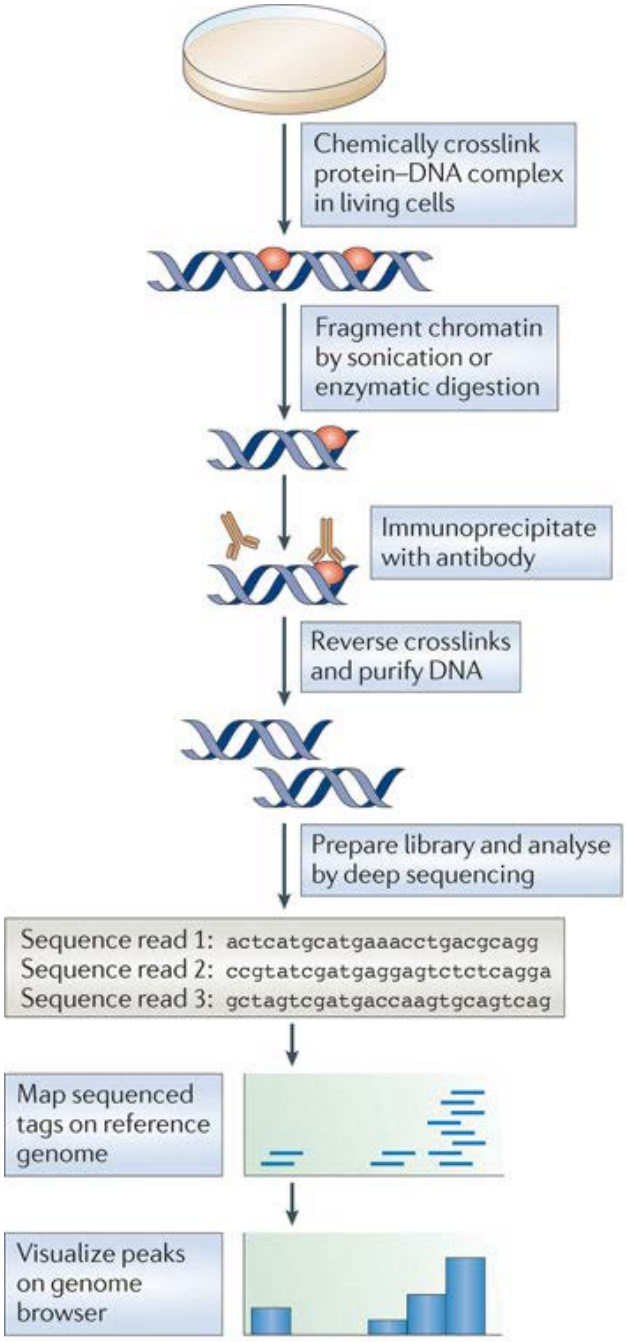


Figure 1.6 ChIP-Seq technique combines chromatin immunoprecipitation (ChIP) with next-generation sequencing to map all protein–DNA interactions genome-wide. *From [81].*

The first is the ChIP, which enriches particular DNA-protein complexes by selecting them with a specific antibody against the protein of interest; the quality of the final results will strongly depend on this first step and, in particular, on the affinity of the antibody for the protein of interest. The second part begins with the purified DNA fragments, which are sequenced simultaneously through a genome analyzer; millions of sequenced reads are then aligned to the reference genome and the IP-enriched regions, or "peaks" – corresponding to the sites where the protein of interest was originally recruited to – are identified.

1.3.2. Data processing issues from sequencing experiments

As every new technology, also ChIP-Seq has certain limitations, many of which arise during the experimental part of the procedure, with a huge impact on the downstream data analysis and therefore on the final results. One example is the amount of starting material. Standard ChIP-Seq protocols typically require large quantities of starting material, in the range of 10^7 cells, in order to obtain a quantity comprised between 1 and 10 ng of immunoprecipitated DNA. Gilfillan and colleagues have reported that a lower cell number risks the involvement of too many cycles of polymerase chain reaction (PCR) to obtain the desired quantity of DNA [82]. The direct consequence of these extra-amplification runs is a drastic reduction of the sensitivity of the method and the rise of PCR duplicates artifacts (that show up as "stacks" of sequences on the genome), which can be wrongly interpreted as binding sites during the peak calling step of the analysis. Conversely, loading too much material on the slide (in the case of Illumina sequencing) can result in fluorescent labels too close to each other to be correctly distinguished, causing lower data quality [83].

Another critical determinant of the quality of the ChIP-Seq experiment is the antibody. Choosing a specific antibody will result in a high level of IP enrichment over the background, while using a non-specific antibody will give rise to false positive peaks. Therefore, apart from testing the enrichment obtained at known binding sites prior the sequencing, it is recommendable to always couple the IP to a control sample, which will ease the peak calling step and avoid misleading enrichment events. Commonly used control samples are: (i) the "input" DNA, that is, DNA prior immunoprecipitation, (ii) the "mock" IP, treated like the IP sample but without any antibody, or (iii) a non-specific antibody, such as anti-IgG.

The choice of using a limited number of sequencing cycles – for example, resulting in DNA sequences (usually called "tags") 25 or 36 basepairs long – can also be problematic during the mapping step of the analysis, as short sequences can be ambiguously mapped to repeated or

low-complexity regions of the reference genome. Moreover, as the number of sequencing cycles increases, there is an exponential decay (or “fading”) in the fluorescent signal intensity, which often results in wrongly called bases [84] and, therefore, in errors during the mapping step of the analysis. Using longer reads (e.g. 50 basepairs long) or paired-end sequencing, in combination with a quality filter of the sequenced tags, can help to reduce this type of problems.

1.3.3. Computational challenges of coregulator ChIP-Seq data analysis

The growing interest in the molecular function of nuclear receptors and coregulator proteins has led to an increase in the number of ChIP-Seq studies to identify their binding sites genome-wide. However, coregulators do not bind DNA directly, but rather exert their function by interacting with several other proteins, such as nuclear receptors and chromatin remodeling complexes. Therefore, this indirect action of coregulators poses additional challenges to the analysis of ChIP-Seq data.

First of all, this technique will detect only indirect protein-DNA interactions, making it difficult to trace back which transcription factor was mediating the interaction of the coregulator with the DNA. Moreover, since coregulators are able to control several distinct factors, it is highly probable that different peaks will correspond to different transcription factor instances, albeit bound to the same coregulator. Hence, a *de novo* motif analysis of the ChIP-Seq peaks will be dominated by the most enriched transcription factor (TF), although other TFs might be bound as well, even if to a minor extent, by the chipped coregulator.

Another challenge in coregulator ChIP-Seq data analysis is a direct consequence of the fact that coregulators can bind to enhancer regions and bring these in contact with promoters – which might be located farther away from the enhancer – to increase the gene transcription rate. As ChIP-Seq blocks the whole DNA-protein complex through formaldehyde, it might happen that an enhancer peak will be wrongly assigned to its nearest gene, rather than to the truly regulated gene, located next to the promoter in contact with the enhancer.

1.4. Genome-wide studies on ERRα and PGC-1α

Based on genetic and molecular approaches in highly oxidative tissues, a number of ERRα and PGC-1α target genes involved in carbohydrate, lipid and mitochondrial metabolism have been identified during the last decade. However, although specific characteristics of both ERRα and

PGC-1α have been thoroughly analyzed by several studies, only a few of them have explored to date the recruitment and the downstream effects of these two proteins on a genome-wide scale.

By performing ChIP-on-chip assay and microarray analysis on wild-type and ERRα null mice hearts, the group of Vincent Giguère showed that ERRα is involved in the uptake, production and transport of energy substrates across mitochondrial membranes, as well as in the intracellular sensing of altered Ca^{2+} levels, which represent the signal of an increased demand of energy sources [85]. Using a similar genome-wide approach, the same group later identified ERRα as a regulator of renal Na^{+}/K^{+} homeostasis and of the renin-angiotensin pathway in mice kidneys [86]. In liver, a ChIP-Seq experiment on endogenous ERRα showed that the mammalian target of rapamycin (mTOR) regulates ERRα activity by repressing its ubiquitin-mediated degradation; in fact, rapamycin treatment, as well as genetic ablation of ERRα, reduced ERRα ability to control TCA cycle and lipid biosynthesis, causing the development of nonalcoholic fatty liver disease [87]. Finally, a ChIP-on-chip study on ERRα in mouse liver showed a significant overlap between its metabolic target genes and the recruitment sites of the homeobox protein prospero-related homeobox 1 (Prox1); moreover, Prox1 was shown to interact with both ERRα and PGC-1α and to inhibit their activity in the context of energy production [88].

Regarding PGC-1α, the first genome wide study was conducted in 2004 by the Spiegelman's group [35]. By combining a microarray of PGC-1α-induced transcriptional profiles with a computational approach to detect *cis*-regulatory motifs in genome-wide promoter regions, the authors identified ERRα and GABPA as key transcription factors for the regulation of the OXPHOS pathway. In particular, the software they developed, MotifADE, identified the two TFs by (i) ranking all array genes expression fold changes, (ii) performing a k-mer search in their promoter regions and (iii) using the Mann-Whitney rank sum statistic to determine whether the distribution of fold changes for the subset of genes containing a given k-mer was significantly different from the genes lacking the k-mer. Another genome-wide study, conducted in the group of Jiandie Lin, identified the BRG1-associated factor 60a (BAF60a) as a novel partner of PGC-1α in the regulation of hepatic lipid metabolism [89]; through a coactivation reporter screen, the authors created a collection of 1146 transcriptional regulators (corresponding to ~48% of all predicted human TFs and cofactors) and tested whether the activity of individual TFs was enhanced in presence or absence of PGC-1α. To our knowledge, the only ChIP-Seq analysis of PGC-1α was performed by Charos and colleagues in human hepatic carcinoma cells (HepG2) treated with forskolin, an inducer of PGC-1α [90]. A *de novo* motif search in the top 250 peaks identified heat shock factor 1 (HSF1), ERRα and the CCAAT/enhancer-binding protein β (CEBPB), although no experimental validation of these predictions was reported; moreover, the

association of regulator-occupancy with gene expression revealed a stronger correlation of PGC-1α with down-regulated than with up-regulated genes, which is in contrast with the well-known coactivatory function of PGC-1α.

In conclusion, despite some studies have started to explore the regulatory networks controlled by ERRα and PGC-1α, the exhaustive set of partner TFs and of the corresponding downstream targets, in particular for the skeletal muscle tissue, is still far from complete.

1.5. Aims of the thesis

The coactivator PGC-1α and the nuclear receptor ERRα have been extensively shown to regulate key biological processes related to energy metabolism in higher eukaryotes. However, a rigorous and complete genome-wide analysis of their DNA recruitment in skeletal muscle and of their downstream effects on directly-controlled pathways has not been described so far. In this thesis, using a combination of several computational techniques for the analysis of high-throughput sequencing data, we aimed to unravel the complex transcriptional network of skeletal muscle and to identify novel functional partners for PGC-1α and ERRα. Moreover, we sought to elucidate to which extent are PGC-1α-mediated effects dependent or not on ERRα and which are the biological pathways they regulate together, directly and indirectly, in skeletal muscle cells. Furthermore, we tried to definitely answer the long-debated question of a possible PGC-1α role as corepressor by looking at the direct and indirect regulatory effects of PGC-1α on the activity of all known transcription factors. In addition, the lack of a well-defined analysis pipeline for the study of coregulator ChIP-Seq data motivated us to optimize every step of the NGS data processing method and to provide a solution to the numerous artifacts of this recently developed technique.

References

1. Biddie, S.C., S. John, and G.L. Hager, *Genome-wide mechanisms of nuclear receptor action.* Trends in endocrinology and metabolism: TEM, 2010. 21(1): p. 3-9.
2. Aranda, A. and A. Pascual, *Nuclear hormone receptors and gene expression.* Physiological reviews, 2001. 81(3): p. 1269-304.
3. Zhang, Z., et al., *Genomic analysis of the nuclear receptor family: new insights into structure, regulation, and evolution from the rat genome.* Genome research, 2004. 14(4): p. 580-90.
4. Lonard, D.M. and B.W. O'Malley, *Nuclear receptor coregulators: modulators of pathology and therapeutic targets.* Nature reviews. Endocrinology, 2012. 8(10): p. 598-604.
5. Fang, Y., et al., *NURBS: a database of experimental and predicted nuclear receptor binding sites of mouse.* Bioinformatics, 2013. 29(2): p. 295-7.
6. Ochsner, S.A., et al., *Research resource: Tissue-specific transcriptomics and cistromics of nuclear receptor signaling: a web research resource.* Molecular endocrinology, 2010. 24(10): p. 2065-9.
7. Gadaleta, R.M. and L. Magnani, *Nuclear receptors and chromatin: an inducible couple.* Journal of molecular endocrinology, 2014. 52(2): p. R137-49.
8. Glass, C.K. and S. Ogawa, *Combinatorial roles of nuclear receptors in inflammation and immunity.* Nature reviews. Immunology, 2006. 6(1): p. 44-55.
9. John, S., et al., *Interaction of the glucocorticoid receptor with the chromatin landscape.* Molecular cell, 2008. 29(5): p. 611-24.
10. Hurtado, A., et al., *FOXA1 is a key determinant of estrogen receptor function and endocrine response.* Nature genetics, 2011. 43(1): p. 27-33.
11. Stratmann, M. and U. Schibler, *Transcription factor loading: please take my place!* Cell, 2011. 146(4): p. 497-9.
12. Magnani, L., J. Eeckhoute, and M. Lupien, *Pioneer factors: directing transcriptional regulators within the chromatin environment.* Trends in genetics : TIG, 2011. 27(11): p. 465-74.
13. Glass, C.K. and M.G. Rosenfeld, *The coregulator exchange in transcriptional functions of nuclear receptors.* Genes & development, 2000. 14(2): p. 121-41.
14. Giguere, V., et al., *Identification of a new class of steroid hormone receptors.* Nature, 1988. 331(6151): p. 91-4.
15. Greschik, H., et al., *Structural and functional evidence for ligand-independent transcriptional activation by the estrogen-related receptor 3.* Molecular cell, 2002. 9(2): p. 303-13.
16. Zhang, Z. and C.T. Teng, *Estrogen receptor-related receptor alpha 1 interacts with coactivator and constitutively activates the estrogen response elements of the human lactoferrin gene.* The Journal of biological chemistry, 2000. 275(27): p. 20837-46.
17. Horard, B. and J.M. Vanacker, *Estrogen receptor-related receptors: orphan receptors desperately seeking a ligand.* Journal of molecular endocrinology, 2003. 31(3): p. 349-57.
18. Busch, B.B., et al., *Identification of a selective inverse agonist for the orphan nuclear receptor estrogen-related receptor alpha.* Journal of medicinal chemistry, 2004. 47(23): p. 5593-6.
19. Razzaque, M.A., et al., *Estrogen receptor-related receptor gamma has an exceptionally broad specificity of DNA sequence recognition.* Gene, 2004. 340(2): p. 275-82.
20. Zhang, Z. and C.T. Teng, *Estrogen receptor alpha and estrogen receptor-related receptor alpha1 compete for binding and coactivator.* Molecular and cellular endocrinology, 2001. 172(1-2): p. 223-33.
21. Martinez, E., F. Givel, and W. Wahli, *A common ancestor DNA motif for invertebrate and vertebrate hormone response elements.* The EMBO journal, 1991. 10(2): p. 263-8.

22. Eichner, L.J. and V. Giguere, *Estrogen related receptors (ERRs): a new dawn in transcriptional control of mitochondrial gene networks.* Mitochondrion, 2011. 11(4): p. 544-52.
23. Bookout, A.L., et al., *Anatomical profiling of nuclear receptor expression reveals a hierarchical transcriptional network.* Cell, 2006. 126(4): p. 789-99.
24. Luo, J., et al., *Placental abnormalities in mouse embryos lacking the orphan nuclear receptor ERR-beta.* Nature, 1997. 388(6644): p. 778-82.
25. Chen, J. and J. Nathans, *Estrogen-related receptor beta/NR3B2 controls epithelial cell fate and endolymph production by the stria vascularis.* Developmental cell, 2007. 13(3): p. 325-37.
26. Onishi, A., et al., *The orphan nuclear hormone receptor ERRbeta controls rod photoreceptor survival.* Proceedings of the National Academy of Sciences of the United States of America, 2010. 107(25): p. 11579-84.
27. Gan, Z., et al., *Nuclear receptor/microRNA circuitry links muscle fiber type to energy metabolism.* The Journal of clinical investigation, 2013. 123(6): p. 2564-75.
28. Alaynick, W.A., et al., *ERRgamma regulates cardiac, gastric, and renal potassium homeostasis.* Molecular endocrinology, 2010. 24(2): p. 299-309.
29. Narkar, V.A., et al., *Exercise and PGC-1alpha-independent synchronization of type I muscle metabolism and vasculature by ERRgamma.* Cell metabolism, 2011. 13(3): p. 283-93.
30. Wende, A.R., et al., *PGC-1alpha coactivates PDK4 gene expression via the orphan nuclear receptor ERRalpha: a mechanism for transcriptional control of muscle glucose metabolism.* Molecular and cellular biology, 2005. 25(24): p. 10684-94.
31. Luo, J., et al., *Reduced fat mass in mice lacking orphan nuclear receptor estrogen-related receptor alpha.* Molecular and cellular biology, 2003. 23(22): p. 7947-56.
32. Murray, J. and J.M. Huss, *Estrogen-related receptor alpha regulates skeletal myocyte differentiation via modulation of the ERK MAP kinase pathway.* American journal of physiology. Cell physiology, 2011. 301(3): p. C630-45.
33. Schreiber, S.N., et al., *The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor estrogen-related receptor alpha (ERRalpha).* The Journal of biological chemistry, 2003. 278(11): p. 9013-8.
34. Soriano, F.X., et al., *Evidence for a mitochondrial regulatory pathway defined by peroxisome proliferator-activated receptor-gamma coactivator-1 alpha, estrogen-related receptor-alpha, and mitofusin 2.* Diabetes, 2006. 55(6): p. 1783-91.
35. Mootha, V.K., et al., *Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6570-5.
36. Willy, P.J., et al., *Regulation of PPARgamma coactivator 1alpha (PGC-1alpha) signaling by an estrogen-related receptor alpha (ERRalpha) ligand.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(24): p. 8912-7.
37. Lonard, D.M. and W. O'Malley B, *Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation.* Molecular cell, 2007. 27(5): p. 691-700.
38. Fan, W., et al., *Road to exercise mimetics: targeting nuclear receptors in skeletal muscle.* Journal of molecular endocrinology, 2013. 51(3): p. T87-T100.
39. Varga-Weisz, P., *ATP-dependent chromatin remodeling factors: nucleosome shufflers with many missions.* Oncogene, 2001. 20(24): p. 3076-85.
40. Yamamoto, H., et al., *NCoR1 is a conserved physiological modulator of muscle mass and oxidative function.* Cell, 2011. 147(4): p. 827-39.
41. Perez-Schindler, J., et al., *The corepressor NCoR1 antagonizes PGC-1alpha and estrogen-related receptor alpha in the regulation of skeletal muscle function and oxidative metabolism.* Molecular and cellular biology, 2012. 32(24): p. 4913-24.

42. Eisele, P.S. and C. Handschin, *Functional crosstalk of PGC-1 coactivators and inflammation in skeletal muscle pathophysiology.* Seminars in immunopathology, 2014. 36(1): p. 27-53.
43. Spiegelman, B.M. and R. Heinrich, *Biological control through regulated transcriptional coactivators.* Cell, 2004. 119(2): p. 157-67.
44. Arany, Z., et al., *The transcriptional coactivator PGC-1beta drives the formation of oxidative type IIX fibers in skeletal muscle.* Cell metabolism, 2007. 5(1): p. 35-46.
45. Rowe, G.C., et al., *PGC-1beta regulates angiogenesis in skeletal muscle.* American journal of physiology. Endocrinology and metabolism, 2011. 301(1): p. E155-63.
46. Lin, J., et al., *Hyperlipidemic effects of dietary saturated fats mediated through PGC-1beta coactivation of SREBP.* Cell, 2005. 120(2): p. 261-73.
47. Ishii, K.A., et al., *Coordination of PGC-1beta and iron uptake in mitochondrial biogenesis and osteoclast activation.* Nature medicine, 2009. 15(3): p. 259-66.
48. Uldry, M., et al., *Complementary action of the PGC-1 coactivators in mitochondrial biogenesis and brown fat differentiation.* Cell metabolism, 2006. 3(5): p. 333-41.
49. Herzig, S., et al., *CREB regulates hepatic gluconeogenesis through the coactivator PGC-1.* Nature, 2001. 413(6852): p. 179-83.
50. Jager, S., et al., *AMP-activated protein kinase (AMPK) action in skeletal muscle via direct phosphorylation of PGC-1alpha.* Proceedings of the National Academy of Sciences of the United States of America, 2007. 104(29): p. 12017-22.
51. Handschin, C., et al., *An autoregulatory loop controls peroxisome proliferator-activated receptor gamma coactivator 1alpha expression in muscle.* Proceedings of the National Academy of Sciences of the United States of America, 2003. 100(12): p. 7111-6.
52. Daitoku, H., et al., *Regulation of PGC-1 promoter activity by protein kinase B and the forkhead transcription factor FKHR.* Diabetes, 2003. 52(3): p. 642-9.
53. Wang, L., et al., *The orphan nuclear receptor SHP regulates PGC-1alpha expression and energy production in brown adipocytes.* Cell metabolism, 2005. 2(4): p. 227-38.
54. Fernandez-Marcos, P.J. and J. Auwerx, *Regulation of PGC-1alpha, a nodal regulator of mitochondrial biogenesis.* The American journal of clinical nutrition, 2011. 93(4): p. 884S-90.
55. Fan, M., et al., *Suppression of mitochondrial respiration through recruitment of p160 myb binding protein to PGC-1alpha: modulation by p38 MAPK.* Genes & development, 2004. 18(3): p. 278-89.
56. Olson, B.L., et al., *SCFCdc4 acts antagonistically to the PGC-1alpha transcriptional coactivator by targeting it for ubiquitin-mediated proteolysis.* Genes & development, 2008. 22(2): p. 252-64.
57. Lerin, C., et al., *GCN5 acetyltransferase complex controls glucose metabolism through transcriptional repression of PGC-1alpha.* Cell metabolism, 2006. 3(6): p. 429-38.
58. Nemoto, S., M.M. Fergusson, and T. Finkel, *SIRT1 functionally interacts with the metabolic regulator and transcriptional coactivator PGC-1{alpha}.* The Journal of biological chemistry, 2005. 280(16): p. 16456-60.
59. Puigserver, P., et al., *A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis.* Cell, 1998. 92(6): p. 829-39.
60. St-Pierre, J., et al., *Suppression of reactive oxygen species and neurodegeneration by the PGC-1 transcriptional coactivators.* Cell, 2006. 127(2): p. 397-408.
61. Egger, A., et al., *PGC-1alpha determines light damage susceptibility of the murine retina.* PloS one, 2012. 7(2): p. e31272.
62. Yoon, J.C., et al., *Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1.* Nature, 2001. 413(6852): p. 131-8.
63. Handschin, C., et al., *Nutritional regulation of hepatic heme biosynthesis and porphyria through PGC-1alpha.* Cell, 2005. 122(4): p. 505-15.

64. Shin, D.J., et al., *PGC-1alpha activates CYP7A1 and bile acid biosynthesis.* The Journal of biological chemistry, 2003. 278(50): p. 50047-52.
65. Handschin, C., *Regulation of skeletal muscle cell plasticity by the peroxisome proliferator-activated receptor gamma coactivator 1alpha.* Journal of receptor and signal transduction research, 2010. 30(6): p. 376-84.
66. Hood, D.A., et al., *Coordination of metabolic plasticity in skeletal muscle.* The Journal of experimental biology, 2006. 209(Pt 12): p. 2265-75.
67. Hock, M.B. and A. Kralli, *Transcriptional control of mitochondrial biogenesis and function.* Annual review of physiology, 2009. 71: p. 177-203.
68. Wu, Z., et al., *Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1.* Cell, 1999. 98(1): p. 115-24.
69. Pilegaard, H., B. Saltin, and P.D. Neufer, *Exercise induces transient transcriptional activation of the PGC-1alpha gene in human skeletal muscle.* The Journal of physiology, 2003. 546(Pt 3): p. 851-8.
70. Summermatter, S., et al., *Peroxisome proliferator-activated receptor {gamma} coactivator 1{alpha} (PGC-1{alpha}) promotes skeletal muscle lipid refueling in vivo by activating de novo lipogenesis and the pentose phosphate pathway.* The Journal of biological chemistry, 2010. 285(43): p. 32793-800.
71. Vega, R.B., J.M. Huss, and D.P. Kelly, *The coactivator PGC-1 cooperates with peroxisome proliferator-activated receptor alpha in transcriptional control of nuclear genes encoding mitochondrial fatty acid oxidation enzymes.* Molecular and cellular biology, 2000. 20(5): p. 1868-76.
72. Lin, J., et al., *Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres.* Nature, 2002. 418(6899): p. 797-801.
73. Terada, S., et al., *Effects of low-intensity prolonged exercise on PGC-1 mRNA expression in rat epitrochlearis muscle.* Biochemical and biophysical research communications, 2002. 296(2): p. 350-4.
74. Handschin, C., et al., *Skeletal muscle fiber-type switching, exercise intolerance, and myopathy in PGC-1alpha muscle-specific knock-out animals.* The Journal of biological chemistry, 2007. 282(41): p. 30014-21.
75. Shendure, J. and H. Ji, *Next-generation DNA sequencing.* Nature biotechnology, 2008. 26(10): p. 1135-45.
76. Desmond-Hellmann, S., *Toward precision medicine: a new social contract?* Science translational medicine, 2012. 4(129): p. 129ed3.
77. Feuk, L., A.R. Carson, and S.W. Scherer, *Structural variation in the human genome.* Nature reviews. Genetics, 2006. 7(2): p. 85-97.
78. Hafner, M., et al., *Transcriptome-wide identification of RNA-binding protein and microRNA target sites by PAR-CLIP.* Cell, 2010. 141(1): p. 129-41.
79. Johnson, D.S., et al., *Genome-wide mapping of in vivo protein-DNA interactions.* Science, 2007. 316(5830): p. 1497-502.
80. Robertson, G., et al., *Genome-wide profiles of STAT1 DNA association using chromatin immunoprecipitation and massively parallel sequencing.* Nature methods, 2007. 4(8): p. 651-7.
81. O'Shea, J.J., et al., *Genomic views of STAT function in CD4+ T helper cell differentiation.* Nature reviews. Immunology, 2011. 11(4): p. 239-50.
82. Gilfillan, G.D., et al., *Limitations and possibilities of low cell number ChIP-seq.* BMC genomics, 2012. 13: p. 645.
83. Park, P.J., *ChIP-seq: advantages and challenges of a maturing technology.* Nature reviews. Genetics, 2009. 10(10): p. 669-80.
84. Erlich, Y., et al., *Alta-Cyclic: a self-optimizing base caller for next-generation sequencing.* Nature methods, 2008. 5(8): p. 679-82.

85. Dufour, C.R., et al., *Genome-wide orchestration of cardiac functions by the orphan nuclear receptors ERRalpha and gamma.* Cell metabolism, 2007. 5(5): p. 345-56.
86. Tremblay, A.M., et al., *Physiological genomics identifies estrogen-related receptor alpha as a regulator of renal sodium and potassium homeostasis and the renin-angiotensin pathway.* Molecular endocrinology, 2010. 24(1): p. 22-32.
87. Chaveroux, C., et al., *Molecular and genetic crosstalks between mTOR and ERRalpha are key determinants of rapamycin-induced nonalcoholic fatty liver.* Cell metabolism, 2013. 17(4): p. 586-98.
88. Charest-Marcotte, A., et al., *The homeobox protein Prox1 is a negative modulator of ERR{alpha}/PGC-1{alpha} bioenergetic functions.* Genes & development, 2010. 24(6): p. 537-42.
89. Li, S., et al., *Genome-wide coactivation analysis of PGC-1alpha identifies BAF60a as a regulator of hepatic lipid metabolism.* Cell metabolism, 2008. 8(2): p. 105-17.
90. Charos, A.E., et al., *A highly integrated and complex PPARGC1A transcription factor binding network in HepG2 cells.* Genome research, 2012. 22(9): p. 1668-79.

2. ChIP-Seq data analysis – From raw data to peaks identification

2.1. ChIP-Seq platforms and technologies

Nowadays, the most commonly used ChIP-Seq platforms are provided by Illumina, Ion Torrent and Roche 454 Pyrosequencing. In the latter technique, each DNA fragment (up to 1 kilobase, or kb) is fixed, through a generic adaptor, to a DNA-capture bead and amplified. Each bead is then placed, together with DNA polymerase and sequencing buffers, into a single well of a slide, where the reaction of synthesis takes place; every time the slide is flooded with one of the four nucleotides, the successful incorporation of the base in the DNA releases a light signal, whose density is detected and used to reconstruct the sequence.

In contrast, Ion Torrent technology does not make use of any optical signal; instead, it determines how many nucleotides (if any) are added at each cycle by measuring the change in pH which happens upon the release of an H+ ion following the addition of a base to the DNA polymer.

The procedure adopted by the Illumina Genome Analyzer is the most alike to the classical Sanger sequencing, as it also exploits terminator bases to stop the PCR reaction as soon as they are incorporated in the growing DNA. However, the biggest novelty of the Illumina technology consists in the fact that these nucleotides are "reversible terminator bases" (RT-bases) and, as such, once the fluorescent dye is chopped by the cleavage enzyme, they revert to normally functioning nucleotides [1]. This property of the reagent allows to speed up the so called process of "sequencing-by-synthesis" and to reach a significantly higher throughput compared to Sanger sequencing (Fig.2.1).

More in detail, the Illumina sequencing process, which is the one we chose to generate most of the data contained in this thesis, begins with the library preparation. During this phase, which takes about 6 hours, the protein of interest is cross-linked by formaldehyde to the DNA and the genomic material is sheared by sonication into fragments that, ideally, should result in a range between 100 and 1000 basepairs (bp) long [2]. This range can be further reduced and become much tighter after the size selection process. The following step consists in the addition of bead-attached antibodies, raised against the protein of interest, to immunoprecipitate specific DNA-

protein complexes; in parallel, a sample of the whole cell extract (WCE, also called "Input DNA") is put aside and later used as the background to determine the enrichment of the target protein. After purification of the DNA-protein complexes, generic adapter sequences are ligated to the ChIP DNA and to the WCE, and both samples undergo a PCR amplification step in order to get enough starting material for the sequencing reaction.

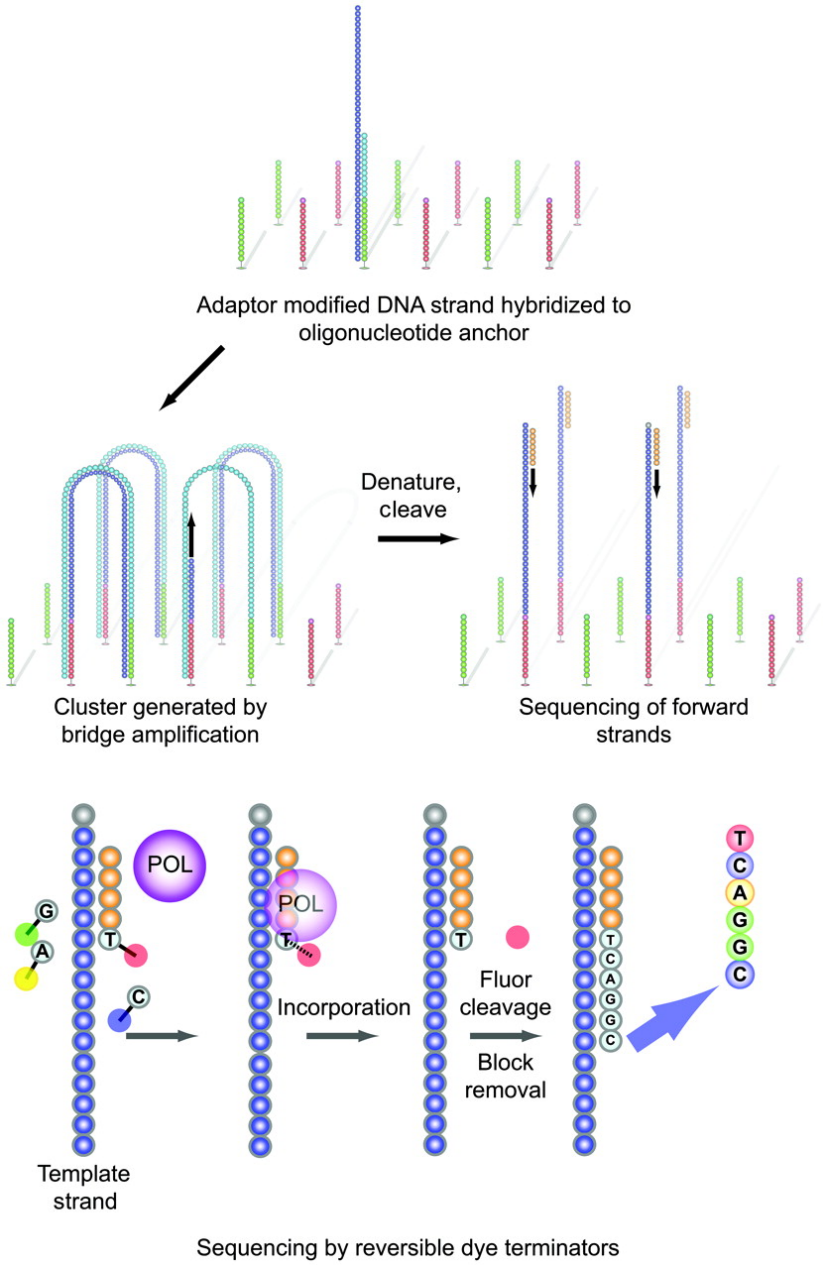


Figure 2.1 The Illumina sequencing method. *From [3]*.

Single strand ChIP DNA and WCE oligos are loaded on individual lanes of an Illumina slide, named "flowcell", and hybridized to it through the previously ligated adapters. Next, DNA oligos

undergo about 5 hours of automated cluster generation. During this process, called "bridge amplification", the DNA templates are repeatedly: (i) denatured, (ii) hybridized (or "bridged") to a complementary primer, (iii) extended through forward and reverse primers which are covalently attached to the flowcell and (iv) linearized again. Through this solid-phase amplification, single DNA molecules are turned into clonal clusters of up to 1000 identical copies, ending with the production of 100-200 millions of spatially separated clusters across the flowcell surface.

To determine the sequence of each of these clusters, the four previously mentioned RT-bases are added to the flowcell so that the DNA polymerase enzymes can start incorporating them in the nascent DNA strands simultaneously. Cycle after cycle, a new nucleotide is added to each synthesized strand (based on the sequence of the template) and the fluorescence emitted by the dye bound to the incorporated base is detected by a camera. The sequential images captured during this sequencing-by-synthesis are processed in the following image analysis step, which includes clusters identification, background correction and finally base calling, through which the fluorescence signals are converted into actual data (the sequenced "reads" or "tags") together with the corresponding quality scores.

2.2. Quality filter of sequenced reads

The most commonly used file format to store raw sequencing reads obtained from an Illumina instrument is the "FASTQ format". Similarly to the classical FASTA format, a FASTQ file is a text-based human readable format, which consists of a set of biological sequences, each preceded and represented by a one-line header. However, the fundamental difference with FASTA files is the possibility to store also reads' corresponding quality scores, as calculated by the sequencing instrument. This new piece of information is listed right after the read's sequence, such that each biological sequence in the FASTQ file is composed of the following four lines:

- line 1 begins with the "@" character, followed by a sequence identifier; for Illumina data, the identifier usually contains: the instrument name, the flowcell lane number, the tile number with 'x'- and 'y'-coordinates, and, in some cases, a "Y" or "N" whether the read failed the instrument's filter or not, respectively;

- line 2 is the sequence of the read, which consists of the symbols "A", "C", "G", "T", "N" according to the IUPAC notation;

- line 3 begins with the "+" character and is an optional identifier;

- line 4 encodes the quality values of the read in ASCII characters and its length correspond to that of line 2.

The quality score Q associated to a given nucleotide is a value related to the probability p that the corresponding base call is incorrect [4]. Since quality scores were introduced, two different equations have been used to define the value of Q. The first, adopted for Sanger sequencing and named after the PHRED software, was defined in terms of the estimated probability of error:

$$Q_{PHRED} = -10log_{10}p$$

To translate quality scores in single characters (or bytes), FASTQ files used ASCII 33-126 to encode PHRED qualities from 0 to 93, thus adopting an offset of 33.

The second equation, introduced by the Solexa pipeline in 2004, substituted the simple probability p with the odds $p/(1-p)$:

$$Q_{SOLEXA} = -10log_{10}\left(\frac{p}{1-p}\right)$$

For this new formula, an offset of 64 was chosen to translate quality scores in ASCII characters from 59 to 126. Values obtained using the two equations above could be mapped to each other through the following conversions, as defined in [5]:

$$Q_{PHRED} = -10log_{10}\left(10^{Q_{SOLEXA}/10} + 1\right)$$

$$Q_{SOLEXA} = -10log_{10}\left(10^{Q_{PHRED}/10} - 1\right)$$

Finally, after Solexa Inc. was acquired by Illumina in 2006, the original PHRED equation was reintroduced; however, to be interchangeable with the earlier Solexa files, an offset of 64 (rather than 33) was adopted. For an overview of the different mappings of quality scores to ASCII characters, see Figure 2.2.

Since Li & Homer showed in 2010 that using a base quality filter halves the frequency of incorrect read alignment errors [6], several research groups are becoming more and more aware of the importance of filtering low-quality reads to improve the accuracy of the downstream analysis steps [7-9].

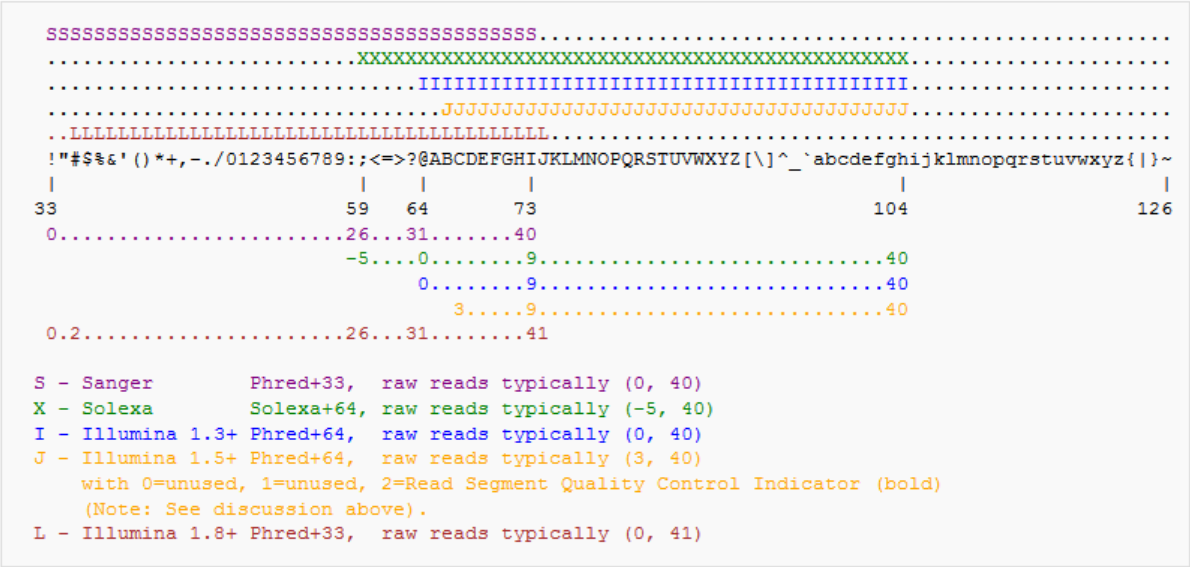


Figure 2.2 Base quality convertion to ASCII characters. *From Wikipedia, retrieved March 25, 2014, from http://en.wikipedia.org/wiki/FASTQ_format.*

Because of the different versions of quality encodings, the most arduous part of the filtering is the determination of the instrument version that was used to sequence the data. Once this obstacle is overcome, the filtering algorithm consists in translating back the ASCII characters of each sequenced read into qualities and deciding whether to keep, discard or trim the read (i.e. cut bases off from the low-quality end) according to specified cutoffs. To filter our ChIP-Seq datasets, we decided to keep only those reads with: (i) a minimum length of 25 basepairs, (ii) an average minimum quality of Q = 20 (corresponding to an error rate of 1%) and (iii) no more than 2 undefined bases (reported in the FASTQ file as "N").

Essentially, the algorithm we developed analyses the first million tags in the input file and collects from these the numerical values corresponding to minimum and maximum ASCII characters, in order to decipher the codification which was used by the instrument. Once this is accomplished, the program parses the file one record (i.e. one read) at the time and scans its quality string symbol by symbol; whenever the average quality falls below the user-defined cutoff, if the length of the read till the current basepair is bigger than the defined length cutoff, the read is trimmed and written in the output file; otherwise, it is skipped. Also, if the DNA sequence contains more than 2 Ns, that read is not reported in the output file.

A useful way to visualize the quality of the sequenced reads before and after the filtering process, is provided by software like FastQC (www.bioinformatics.babraham.ac.uk/projects/fastqc), which provide a set of useful quality

checks (like "Basic Statistics", "Per base GC content" or "Sequence Duplication Levels") on raw sequence data coming from high throughput experiments. To visualize and confirm the results of our quality filter, we looked in particular at the "Per base sequence quality" feature, as shown in Figure 2.3.

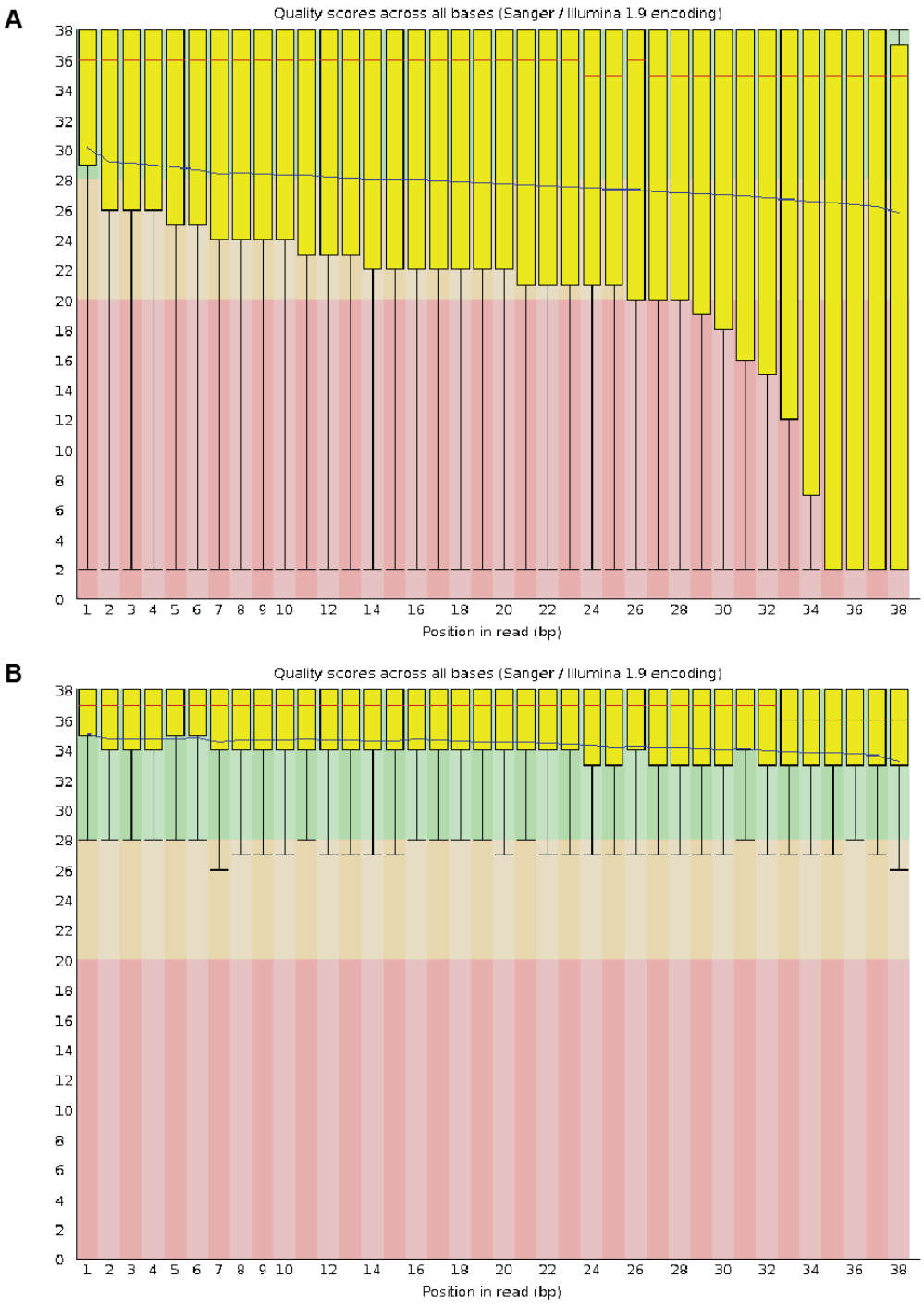


Figure 2.3 Bar plot showing the quality scores of reads of the second IP replicate from the PGC-1alpha sequencing before (**A**) and after (**B**) the quality filter. The y axis is the quality score; colored regions

represent low (red), medium (orange) and high (green) qualities. Each bar represents the quantiles and average quality of the first to the nth base over all reads in the input file. It can be noticed that, as usual, quality degrades toward the end of the sequence.

2.3. Mapping to the reference genome

The next step after the quality filter is the mapping of the sequenced reads to the reference genome. Nowadays, there are several programs to performs this task; in case of reads obtained from a Solexa/Illumina instrument, the currently most used tools are MAQ [10], SOAP [11], Bowtie [12] and BWA [13]. Bowtie in particular, which we chose to align reads from our datasets, has the advantage of being particularly fast, space efficient and at the same time very accurate, thanks to the implementation of an algorithm called "Burrows-Wheeler transform". This method, originally developed for data compression, rearranges a character string by sorting all lexicographic rotations of the original text, then sorting all rows in alphabetical order and finally extracting from this matrix the last column. In this way, it is possible to create an index of a whole mouse genome that occupies only about 2.4 GB on disk and that requires less than 2 GB of RAM for the alignment.

Bowtie has many command-line options that can be tuned according to the desired result; the most commonly used options are described in [14]. For our purpose, we decided to avoid the common practice of discarding multi-mappers, as it leads to a significant loss of information in case, for example, the reads map to duplicated regions of the genome or the protein of interest binds frequently to repeated DNA sequences. For this reason, we used the combination of parameters "--best --strata -a -m 100" to report all valid alignments per read which were "best" in terms of stratum (i.e. number of mismatches) and in terms of the mismatched position(s) quality values; moreover, the option "-m" suppressed all reads which had more than 100 reportable alignments. In the following steps, for those tags mapped to more than one position (i.e. multi-mappers) we assigned to each locus a weight corresponding to the fraction of all alignments reported for that particular read, that is 1/number of positions the read maps to.

Table 2.1 shows the percentages of reads, from our PGC-1alpha datasets, which were successfully mapped using Bowtie. It is worth noting that, in general, for mouse ChIP-Seq data a percentage of uniquely mapped reads higher than 70% is considered normal, while a percentage lower than 50% might indicate problems with the sequencing platform or that the chipped protein might bind frequently in repetitive DNA regions [4].

Table 1 Bowtie report for the 4 samples of the PGC-1α sequencing.

Reads subset	IP 1st repl.	WCE 1st repl.	IP 2nd repl.	WCE 2nd repl.
Processed	6'711'717	17'899'074	36'580'431	35'525'221
With at least one reported alignment	5'699'648 (84.92%)	16'053'370 (89.69%)	21'448'059 (58.63%)	32'244'584 (90.77%)
Failed to align	628'807 (9.37%)	755'365 (4.22%)	13'863'112 (37.90%)	1'140'329 (3.21%)
With alignments suppressed due to -m	383'262 (5.71%)	1'090'339 (6.09%)	1'269'260 (3.47%)	2'140'308 (6.02%)

2.4. Removal of ChIP-Seq artifacts

Right after the mapping step, the mappings can be transformed into genomic densities and visualized in a genome browser, like for example the University of California at Santa Cruz (UCSC) Genome Browser [15] or the Integrative Genomics Viewer (IGV) from the Broad Institute [16]. Specifically, the file containing the mapped alignments per read, usually stored in a "SAM" or "BAM" (i.e. the binary version of SAM) format [17], need to be converted to a "WIG" or "BigWIG" file [15], which contains the number of reads per position at single basepair resolution.

Viewing the data in the context of its genomic region can turn out to be quite informative; in fact, a visual inspection of the data can help to identify particular artifacts like Kb-large regions with an enrichment that is many orders of magnitude higher than the surrounding density, or like "stacks" of reads (Fig.2.4).

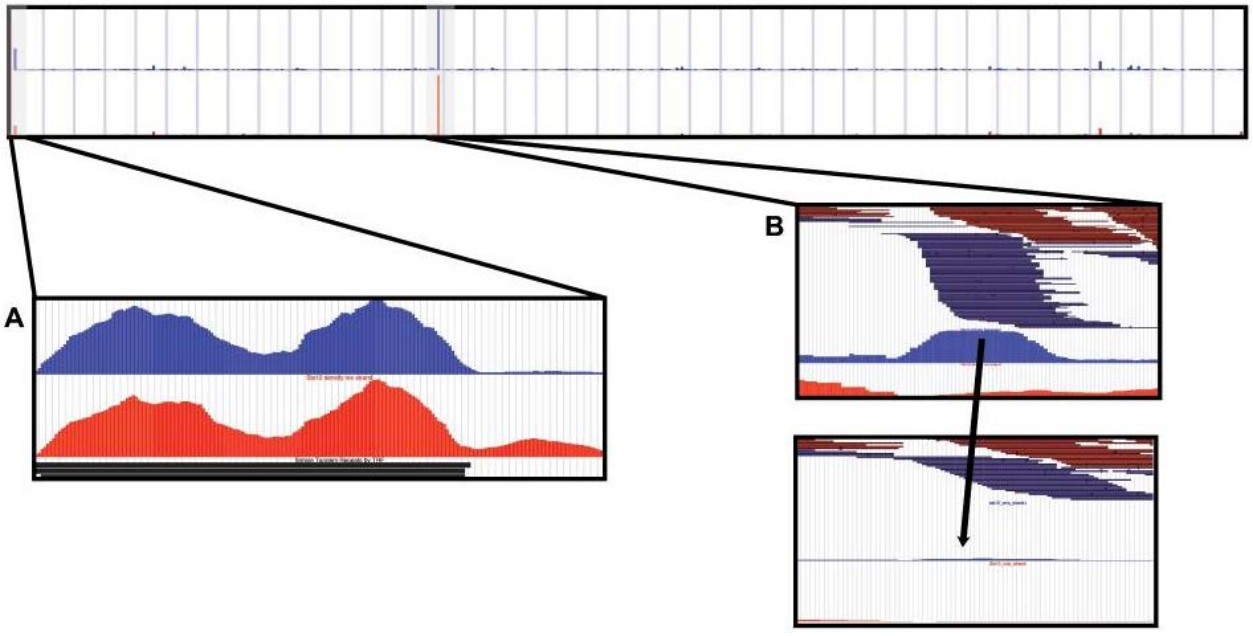


Figure 2.4 Most common ChIP-Seq artifacts. The upper panel depicts an entire chromosome where two regions stand out as highly sampled. (**A**) False enrichment caused by repeated or low-complexity regions (shown by the thick black lines). Red and blue represent reads on forward and reverse strand,

respectively. (**B**) Probable PCR amplification error causing a “stack” of reads. Dark blue lines represent the single reads mapped on the genome and piled on each other, producing the profile in light blue. Upper and lower panel illustrate the “stack” before and after duplicates removal and shows that the light blue profile disappears. *From [18].*

For the first case (schematized by panel A in figure 2.4, from [18]), the principal sources of error might be mapping mistakes around low-complexity regions or genomic repeated locations whose length have been underestimated at the moment of the assembly, while the second artifact (schematized by panel B in figure 2.4, from [18]) originates mainly from PCR amplification errors arising during the sample preparation step.

To disentangle the first type of artifact, we used a 2 kb long sliding window and counted the amount of control reads (from the WCE sample) genome-wide; the distribution of these counts (Fig.2.5) helped us to set a sample-specific cutoff to the maximum number of background reads which was used in the following peak finding step. This way, we prevented calling as (false positive) peaks enriched regions located in repeated or low-complexity regions (Fig.2.6).

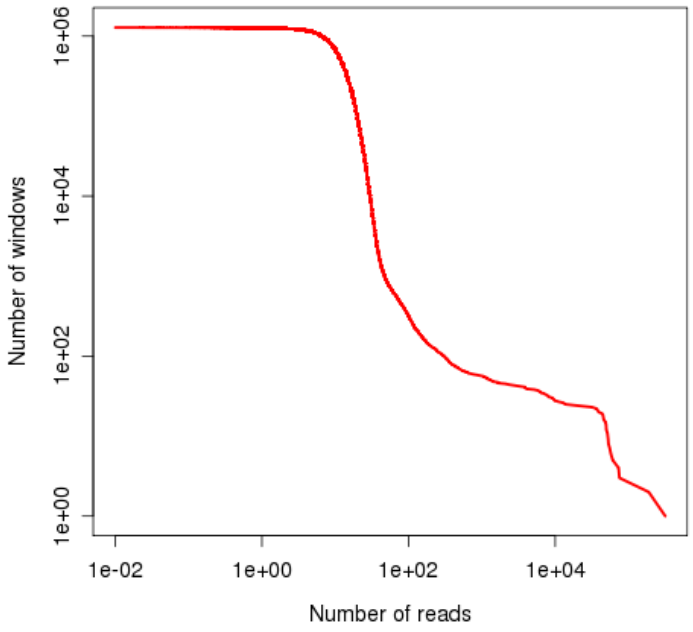


Figure 2.5 Control reads distribution in the first replicate sample from the PGC-1α sequencing.

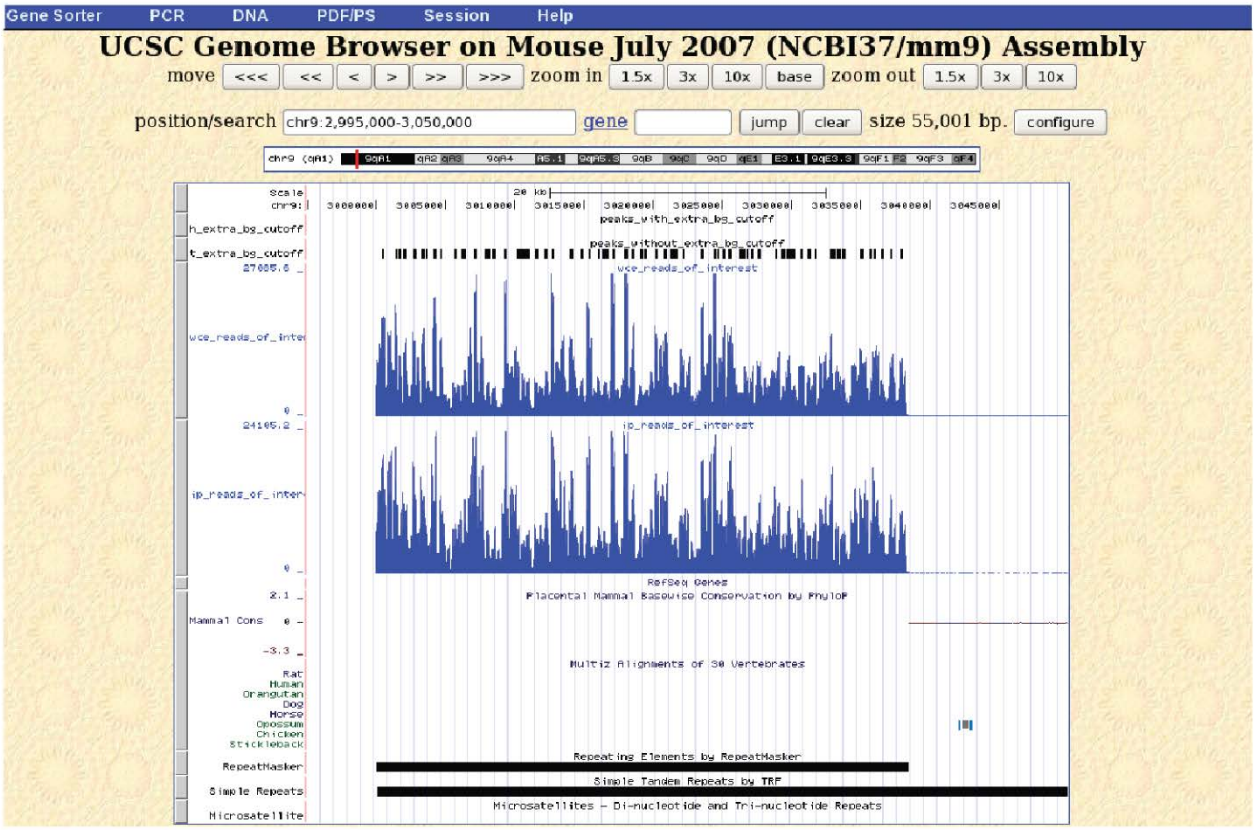


Figure 2.6 Example of a set of false positive peaks (blue distributions) in a region overlapping repeated DNA. Read distributions come from the PGC-1α sequencing data (upper track: WCE, lower track: IP).

Our solution to the second type of artifact was to assign a weight of 1 to any pile of tags starting at the same genomic position, as if that locus was covered just by one read. The rationale for this procedure is that it would be very unlikely for the sonication step to break DNA at the exactly same location more than twice by chance; hence, if we see more than one read per locus, it was most probably caused during the PCR amplification step.

An example of a PCR amplification error next to a bona fide tag enrichment is illustrated in Figure 2.7, while Figure 2.8 shows the distribution of the IP reads from the ERRα sequencing experiment, in which it can be clearly noticed that most of the loci in the genome were covered by around 100 reads, corresponding to the average height of the observed stacks.

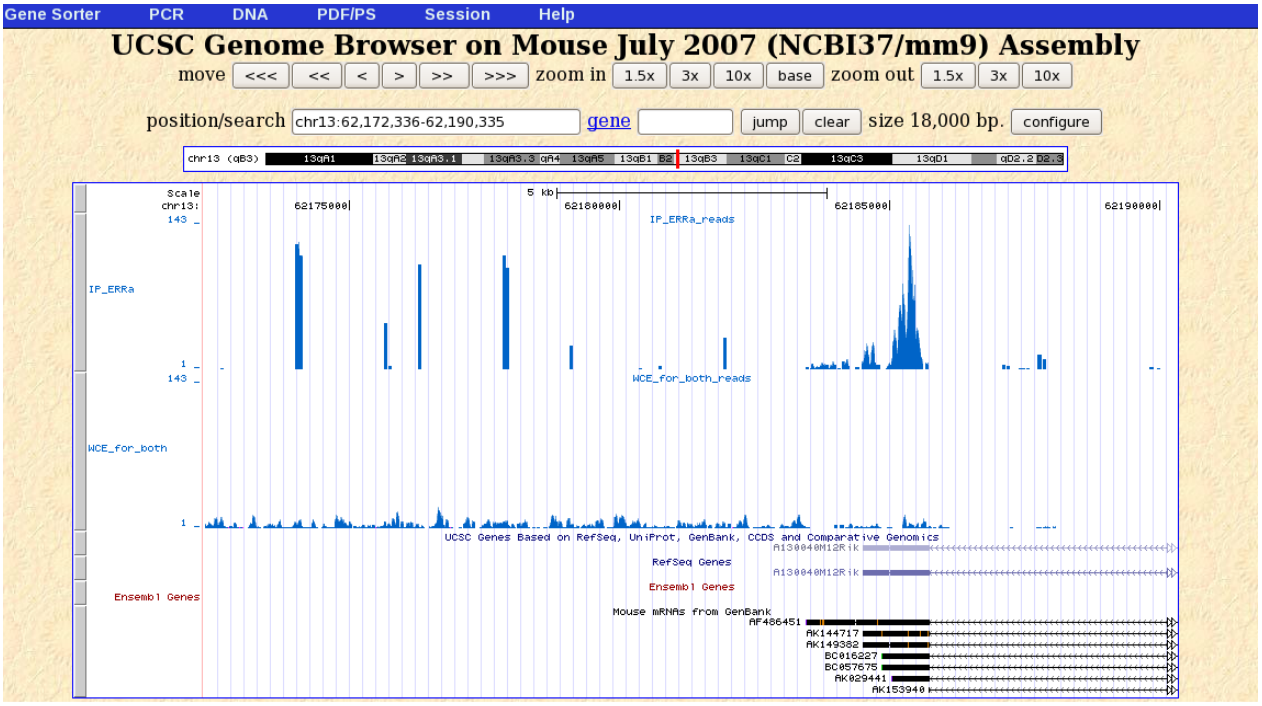


Figure 2.7 Example of "stacks" of reads (blue vertical lines) next to a real peak (in the right portion of the picture). Read distributions come from the ERRα sequencing data (upper track: IP, lower track: WCE).

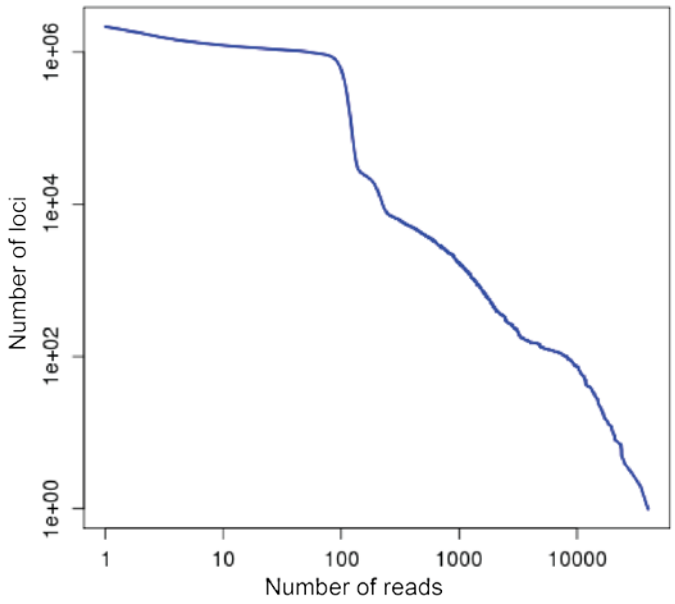


Figure 2.8 Reverse cumulative distribution of the number of reads per loci in the IP sample from the ERRα sequencing experiment.

2.5. Fragment size estimation and read shift

Typically, ChIP-Seq library preparation protocols require a selection step to choose the most appropriate DNA fragment size to use for cluster generation. In the Illumina platforms, only one end of a fragment is read by the instrument and the final size of the sequenced tag is in general much smaller than the one of the original fragment. As a consequence, the distribution of tags on the genome will show two enrichment peaks (on opposite strands) in the surroundings of each actual recruitment site of the chipped protein, with the distance between these peaks being about half of the selected fragment size (Fig.2.9).

A precise identification of the specific sequenced fragment size is thus essential to accurately determine the binding site locations of the protein of interest. However, although recent technologies like the Agilent Bioanalyzer provide a highly precise sample sizing and quantitation, a still commonly used method for size selection is to manually cut out from an electrophoresis gel a band of genomic material which is approximately in the desired size range. Moreover, the estimated fragment size is often not known to the bioinformatician and it needs to be inferred from the data.

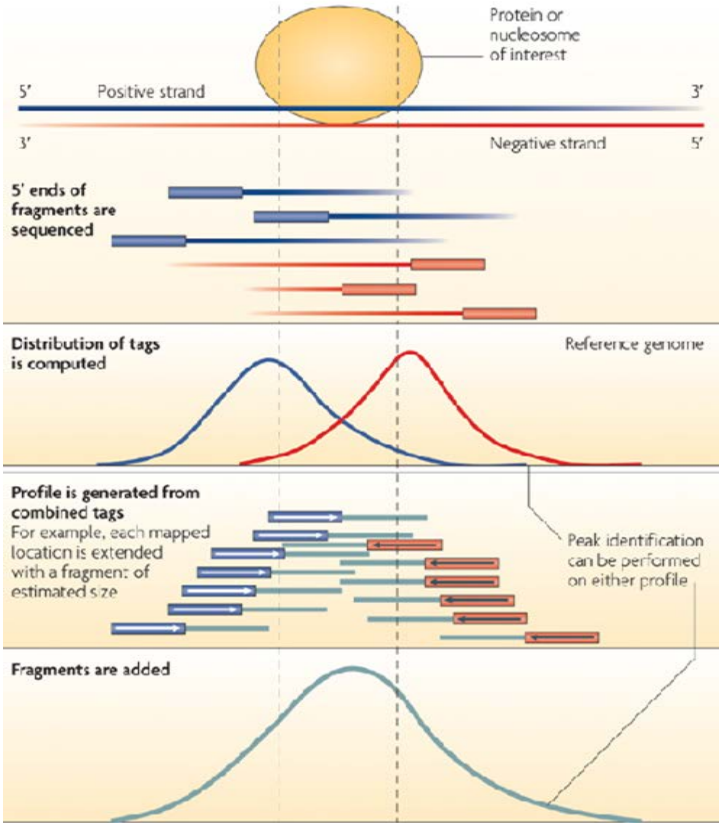


Figure 2.9 Tags distribution around an enriched site of a ChIP-Seq experiment result in two peaks on opposite strands flanking the actual binding location of the protein of interest. *Adapted from [19].*

Our approach to tackle this problem was to find, for each dataset, the maximum value of the cross-correlation $C(d)$ between reads on opposite strands, calculated as:

$$C(d) = \sum_i read_+(i) * read_-(i+d)$$

where $read_+(i)$ and $read_-(i+d)$ are indicator functions that equal 1 if a read starts at position i on the plus strand, or if a read starts at position $i+d$ on the minus strand, respectively. The cross-correlation $C(d)$ was computed for a range of distances d between 1 and 500 bp.

Accordingly, all reads belonging to a given dataset were shifted towards the 3' direction by half of the estimated fragment length, $d/2$, to better represent the precise location of the actual protein recruitment site. As an example, Fig.2.10 shows the cross-correlation or the second replicate IP from the PGC-1α sequencing experiment, whose reads were shifted by 70 bp.

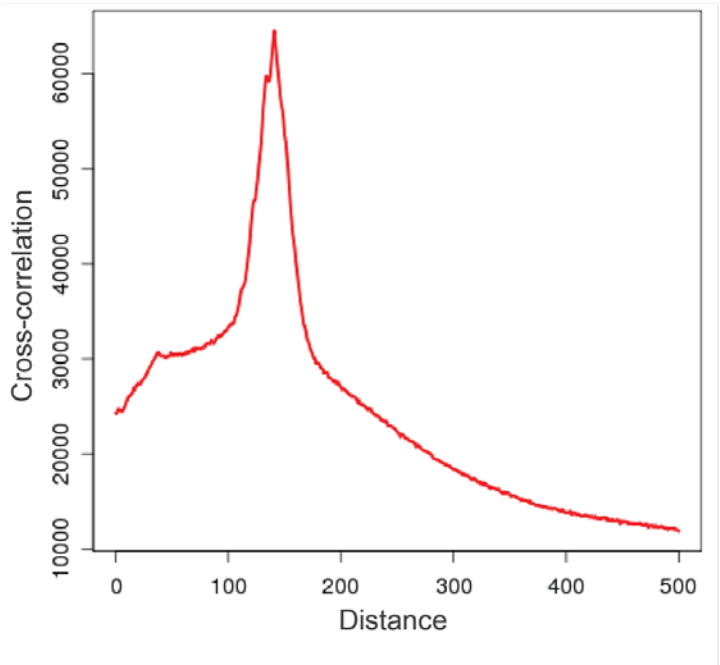


Figure 2.10 Estimation of the average fragment size, measured as the cross-correlation between reads mapped on the plus and minus strands, as a function of their distance. The data is referred to the second replicate IP dataset from the PGC-1α sequencing and the maximum cross-correlation (i.e. the peak of the distribution) was found at a distance of 141 bp.

2.6. Peak calling

A crucial step in ChIP-Seq analysis is to identify the regions in the genome (also called “peaks”) where the protein of interest was significantly enriched. The availability of a sequenced control sample (in our case the whole cell extract or WCE) might help to decrease the number of false positives and, thus, to improve the accuracy of the prediction, although this is not an easy task to prove experimentally on a genome-wide scale. However, when comparing the IP sample to the control dataset, it is important to consider that the total number of reads can differ between samples and, hence, a normalization method must be applied.

In our approach, to make the reads from different samples comparable, we estimate the fraction f_{IP} of all IP reads that are contained in each 200 bp window genome-wide as:

$$f_{IP} = \frac{n_{IP}}{N_{IP}}$$

and, similarly, the fraction f_{WCE} of all WCE reads that are contained in each 2000 bp window, centered on the ChIP one, as:

$$f_{WCE} = \frac{n_{WCE}}{N_{WCE}}$$

where n_{IP} and n_{WCE} are the number of reads in the window from IP and WCE, respectively, while N_{IP} and N_{WCE} are the total number of reads in the IP and WCE samples, correspondingly. In particular, we use a wider window for the control sample to take into account the influence of local fluctuations and we model the tag sampling noise along the genome by a Poisson distribution. The ChIP enrichment of each consecutive window is calculated as:

$$Z = \frac{f_{IP} - f_{WCE}}{\sqrt{{\sigma^2}_{IP} + {\sigma^2}_{WCE}}}$$

where ${\sigma^2}_{IP}$ and ${\sigma^2}_{WCE}$ represent the IP and WCE read frequency variances, given by:

$${\sigma^2}_{IP} = \frac{f_{IP} * (1 - f_{IP})}{N_{IP}} \text{ and } {\sigma^2}_{WCE} = \frac{f_{WCE} * (1 - f_{WCE})}{N_{WCE}}$$

respectively. To speed up the algorithm, we use sliding windows and we move them in steps of 25 bp (Fig.2.11). All consecutive windows having a Z score greater than the chosen cutoff are merged in clusters and the top scoring one (the summit) is called as a peak and used in the following steps.

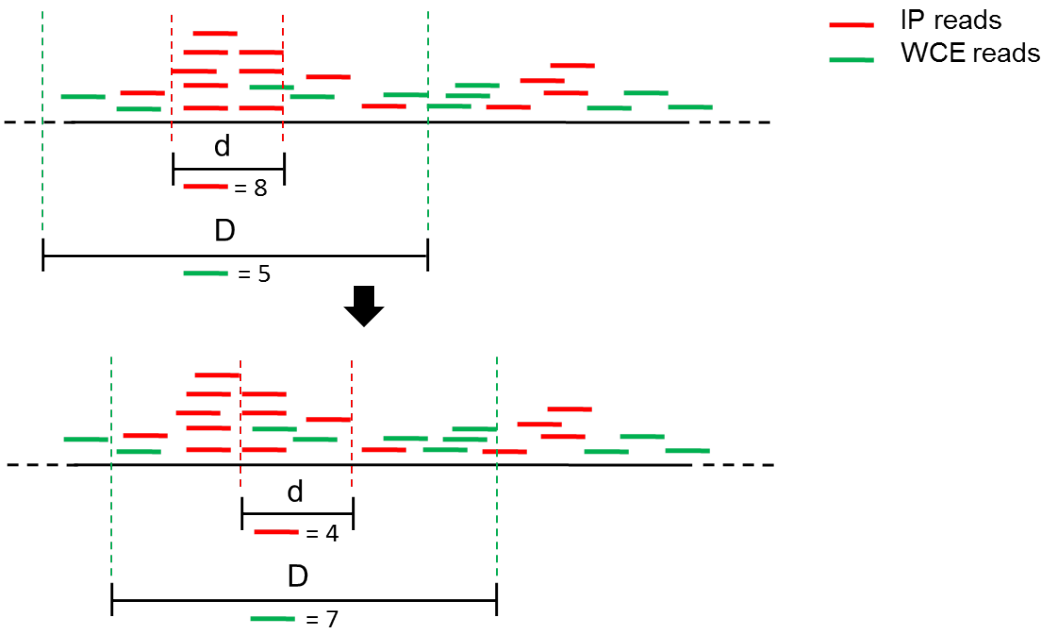


Figure 2.11 Schematic representation of the principle behind the sliding window approach used in the peak finder. “d” and “D” indicate the IP and the WCE sliding windows, respectively.

For the case in which the ChIP-Seq experiment was repeated in biological duplicate, prior to the peak calling step, it is worth looking at the correlation between replicates to evaluate the reproducibility of the ChIP enrichment. One way to proceed is to identify the top peaks in one dataset (using the above formula), to examine the correlation between their Z scores and the Z values of the corresponding regions in the other dataset and to repeat this procedure switching the two duplicates. Figure 2.12 shows the correlation between the first and the second biological replicate of the PGC-1α sequencing.

Once evaluated the correlation between duplicates, the reads can be pulled together to identify the final set of peaks. To this purpose, we defined the combined Z score for the two biological replicates as:

$$Z = \frac{f_{IP_1} + f_{IP_2} - f_{WCE_1} - f_{WCE_2}}{\sqrt{\sigma^2_{IP_1} + \sigma^2_{IP_2} + \sigma^2_{WCE_1} + \sigma^2_{WCE_2}}}.$$

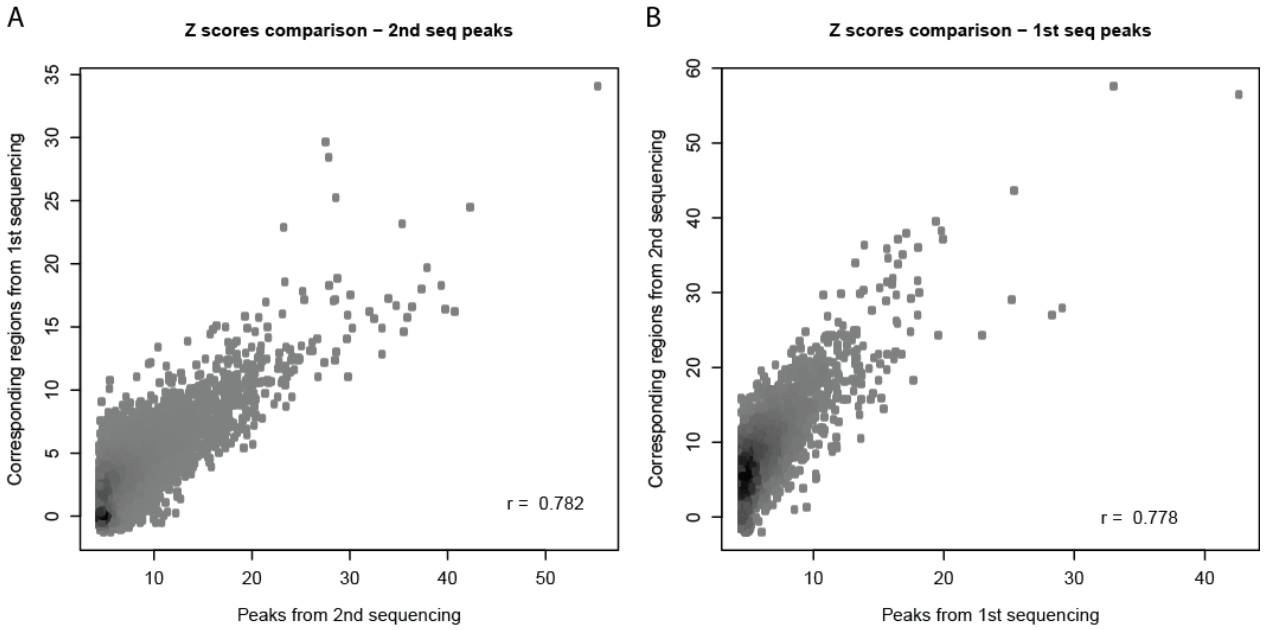


Figure 2.12 The ChIP enrichment was reproducible across biological replicates from the PGC-1α sequencing dataset. (**A**) Correspondence between the Z scores of the peaks called for the first sequencing experiment and the Z scores of the corresponding regions in the second sequencing and (**B**) viceversa. Colors indicate the density of the points (grey:low, black:high). "r" indicates the Pearson correlation coefficient.

References

1. Turcatti, G., et al., *A new class of cleavable fluorescent nucleotides: synthesis and optimization as reversible terminators for DNA sequencing by synthesis.* Nucleic acids research, 2008. 36(4): p. e25.
2. Raha, D., M. Hong, and M. Snyder, *ChIP-Seq: a method for global identification of regulatory elements in the genome.* Current protocols in molecular biology / edited by Frederick M. Ausubel ... [et al.], 2010. Chapter 21: p. Unit 21 19 1-14.
3. Voelkerding, K.V., S.A. Dames, and J.D. Durtschi, *Next-generation sequencing: from basic research to diagnostics.* Clinical chemistry, 2009. 55(4): p. 641-58.
4. Bailey, T., et al., *Practical guidelines for the comprehensive analysis of ChIP-seq data.* PLoS computational biology, 2013. 9(11): p. e1003326.
5. Cock, P.J., et al., *The Sanger FASTQ file format for sequences with quality scores, and the Solexa/Illumina FASTQ variants.* Nucleic acids research, 2010. 38(6): p. 1767-71.
6. Li, H. and N. Homer, *A survey of sequence alignment algorithms for next-generation sequencing.* Briefings in bioinformatics, 2010. 11(5): p. 473-83.
7. Ramos, R.T., et al., *Analysis of quality raw data of second generation sequencers with Quality Assessment Software.* BMC research notes, 2011. 4: p. 130.
8. Mendoza-Parra, M.A., et al., *A quality control system for profiles obtained by ChIP sequencing.* Nucleic acids research, 2013. 41(21): p. e196.
9. Planet, E., et al., *htSeqTools: high-throughput sequencing quality control, processing and visualization in R.* Bioinformatics, 2012. 28(4): p. 589-90.
10. Li, H., J. Ruan, and R. Durbin, *Mapping short DNA sequencing reads and calling variants using mapping quality scores.* Genome research, 2008. 18(11): p. 1851-8.
11. Li, R., et al., *SOAP: short oligonucleotide alignment program.* Bioinformatics, 2008. 24(5): p. 713-4.
12. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome biology, 2009. 10(3): p. R25.
13. Li, H. and R. Durbin, *Fast and accurate short read alignment with Burrows-Wheeler transform.* Bioinformatics, 2009. 25(14): p. 1754-60.
14. Langmead, B., *Aligning short sequencing reads with Bowtie.* Current protocols in bioinformatics / editoral board, Andreas D. Baxevanis ... [et al.], 2010. Chapter 11: p. Unit 11 7.
15. Karolchik, D., et al., *The UCSC Genome Browser database: 2014 update.* Nucleic acids research, 2014. 42(Database issue): p. D764-70.
16. Thorvaldsdottir, H., J.T. Robinson, and J.P. Mesirov, *Integrative Genomics Viewer (IGV): high-performance genomics data visualization and exploration.* Briefings in bioinformatics, 2013. 14(2): p. 178-92.
17. Li, H., et al., *The Sequence Alignment/Map format and SAMtools.* Bioinformatics, 2009. 25(16): p. 2078-9.
18. Leleu, M., G. Lefebvre, and J. Rougemont, *Processing and analyzing ChIP-seq data: from short reads to regulatory interactions.* Briefings in functional genomics, 2010. 9(5-6): p. 466-76.
19. Park, P.J., *ChIP-seq: advantages and challenges of a maturing technology.* Nature reviews. Genetics, 2009. 10(10): p. 669-80.

3. ChIP-Seq data analysis – From peaks to the downstream analysis

3.1. Peak annotation to genomic elements

The major goals of a ChIP-Seq experiment are to determine all locations genome-wide where the protein of interest is recruited to and to relate these binding positions to one or more cellular functions in which the protein might be involved. To address these objectives, ChIP-Seq peaks are usually “annotated” to functionally relevant genomic elements, which include transcripts, promoters, enhancers, intergenic regions, miRNAs, etc. In the absence of a more detailed chromatin conformation caption experiment – which can identify chromosomal regions folded to bring interacting regulatory regions together, for example enhancers and genes – it is common practice to associate ChIP-Seq peaks to their nearest genomic annotation, assuming that the closer the distance, the more likely the association is correct.

Accordingly, we annotated the peaks from our datasets to their closest RefSeq Mus musculus (mouse) transcripts [1] and we defined binding regions as “Intergenic”, “Intronic”, “Exonic”, “Upstream of TSS” or “Downstream of TSS” in case they were, respectively, further than 10 kb from the nearest RefSeq transcript, inside an intron, inside an exon, comprised between a distance of -10 to 0 kb from the nearest transcription start site (TSS), or located between a distance of 0 to 10 kb from the closest transcription end site (TES). In addition, to compare the observed peak location distributions with the expected ones, we generated 100 random peak sets of the same length, checked their positions in relation to the RefSeq transcripts and averaged the results. As shown by Figure 3.1, PGC-1α binding peaks occur twice as often in the proximal promoter region and about 1.3 times more frequently inside the gene body, compared to random peaks. These numbers are even higher in ERRα peaks, which are located more often in the region upstream of TSS (22%) and less frequently in intergenic regions (41%) than the PGC-1α peaks.

In parallel to the above described annotation procedure, which helped us to get an insight about the genomic distribution of PGC-1α and ERRα peaks, both peak sets were also assigned by proximity to the closest promoter regions, which were downloaded from the SwissRegulon database [2]. Each peak was assigned to the nearest promoter (and, thus, to the associated gene) in case the distance between the centers of the two elements was smaller than 10 kb.

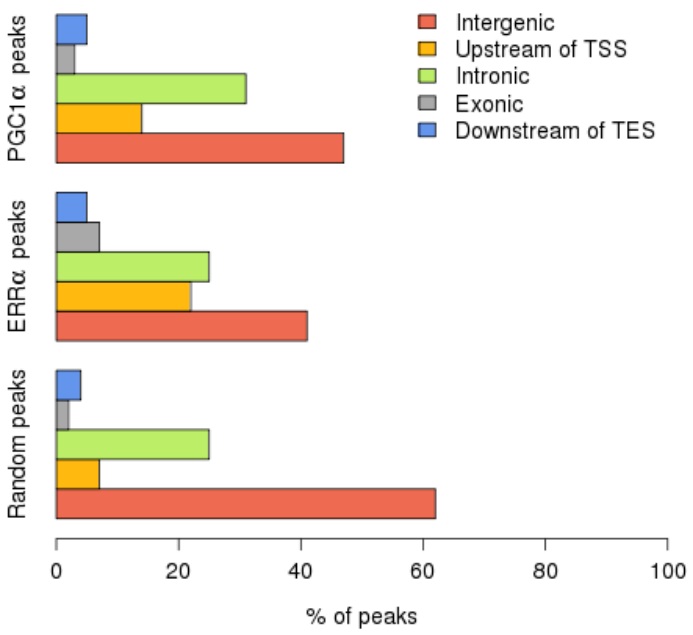


Figure 3.1 Comparison of the percentages of PGC-1α, ERRα and random peaks locations with respect to the nearest RefSeq transcript.

Moreover, taking advantage of the availability of an expression array for PGC-1α over-expression in C2C12 muscle cells, we could make a functional link between the occupancy of PGC-1α along the genome and the regulatory effect on its downstream target genes. In fact, we found that out of the ~30% of our peaks associated to a promoter, 1720 were assigned to a non-changing promoter, 532 to a significantly up-regulated promoter (log_2 foldchange ≥ 2, z-score ≥ 3) and only 43 to a significantly down-regulated promoter (log_2 foldchange ≤ -2, z-score≤ -3), as shown in Figure 3.2.

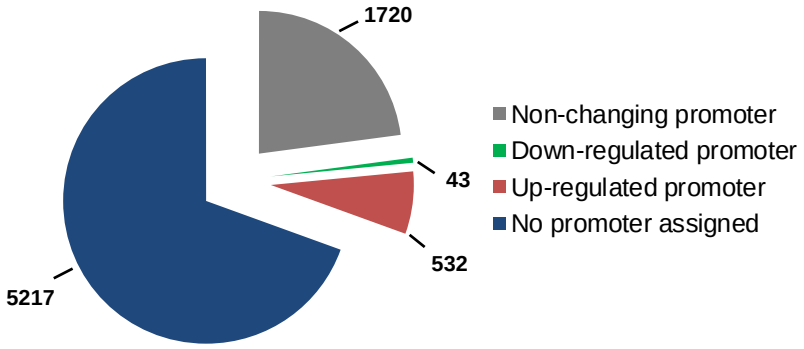


Figure 3.2 Pie chart depicting the number of PGC-1α peaks associated to a mouse promoter (grey, green, red) or not associated to any promoter within the distance of 10 kb (blue).

3.2. Peak co-occurrence with histone modifications

Intriguingly, the vast majority of PGC-1α peaks was not assigned to any promoter within the distance of 10 kb. Our interpretation of these results is that, considering its role as a coregulator, PGC-1α might be recruited to enhancer regions and, thus, be involved in long-range interactions with its downstream target genes. A key limitation of our peak-to-gene assignment, in fact, is that this functional link was restricted to the nearest annotation; although promoter regions are likely to contain a significant number of regulatory sites, it is well known that an important fraction of these, estimated in hundreds of thousands by the ENCODE Project Consortium [3], reside in enhancer regions, also called *cis*-regulatory modules (CRMs). Albeit the experimental identification of these regions is a challenging task, recent developments in high-throughput sequencing have proved that distal regulatory regions can also be recognized by a series of other genomic features, including the methylation pattern, the DNase I hypersensitivity and certain histone modification marks [4-6].

Thanks to the availability of high-throughput data for these genomic traits, we could investigate to which extent our distal peaks were co-occurring with histone modifications by measuring the overlap between the two pieces of information. We downloaded from the GEO database (http://www.ncbi.nlm.nih.gov/geo/) several ChIP-Seq datasets from the Dynlacht lab for histone mark sequencing experiments conducted in C2C12 muscle cells (GEO accession numbers: GSM721289, GSM721291, GSM721293, GSM721295, GSM721297, GSM721301, GSM721303, GSM721305). Although the experimental conditions were not exactly the same as the ones we used for our experiment (in particular the lack of adenoviral infection), the cells used for the histone mark sequencing were at the myotube differentiation stage.

Intriguingly, as shown in Table 3.1, a large fraction of the PGC-1α peaks was indeed overlaying one or more histone marks for open chromatin and active enhancers, strongly suggesting a distal coactivatory role of this protein in the context of skeletal muscle cells. In particular, the histone modifications which PGC-1α peaks were most overlapping to were H3K4me1 and H3K18ac, both found to be present at both active and inactive/poised enhancers [7]. Although H3K4me1 is the most abundant modification among those that were analyzed, the fact that we obtained about half of PGC-1α peaks overlapping a histone marks that occupies only 7.6% of the genome, is most likely not by chance. Similar considerations can be made for the other enhancer marker, H3K18ac, which covers only 3.3% of the genome but is overlapping ~33% of our peaks.

Table 3.1 Overview of the histone modifications for which we evaluated the overlap with the 5217 PGC-1α peaks that were not assigned to any mouse promoter. The histone mark genome coverage was calculated as the ratio between the number of bp covered by the histone mark and the total length of the genome.

Histone modification	Features and functions	PGC-1α peaks overlapping with the histone mark	Histone mark genome coverage	Ratio (fold)
H3K4me1	Marker of for active/poised enhancers at distal loci.	48.5%	7.6%	6.4
H3K4me2	Markers for active/poised promoters. Stimulate the recruitment of chromatin remodeling factors and histone acetyltransferases.	15.4%	1.8%	8.6
H3K4me3		6.5%	2.0%	3.3
H3K27me3	Marker associated with repressed regions and promoters.	4.9%	3.4%	1.4
H3K36me3	Marker for actively transcribed regions and transcriptional elongation.	4.0%	2.4%	1.7
H3K9ac	Marker of active chromatin at promoters and enhancers.	8.3%	0.6%	13.3
H3K18ac	Marker for functional enhancers. Has been linked to activation of nuclear hormone receptors.	33.6%	3.3%	10.2
H4K12ac	Marker enriched at the TSS and along gene bodies. Required for unperturbed progression to S phase.	11.3%	1.0%	11.3

An example of a PGC-1α peak overlapping some histone modification from the ENCODE Project, performed on C2C12 cells at the myoblast stage, is shown in Figure 3.3.

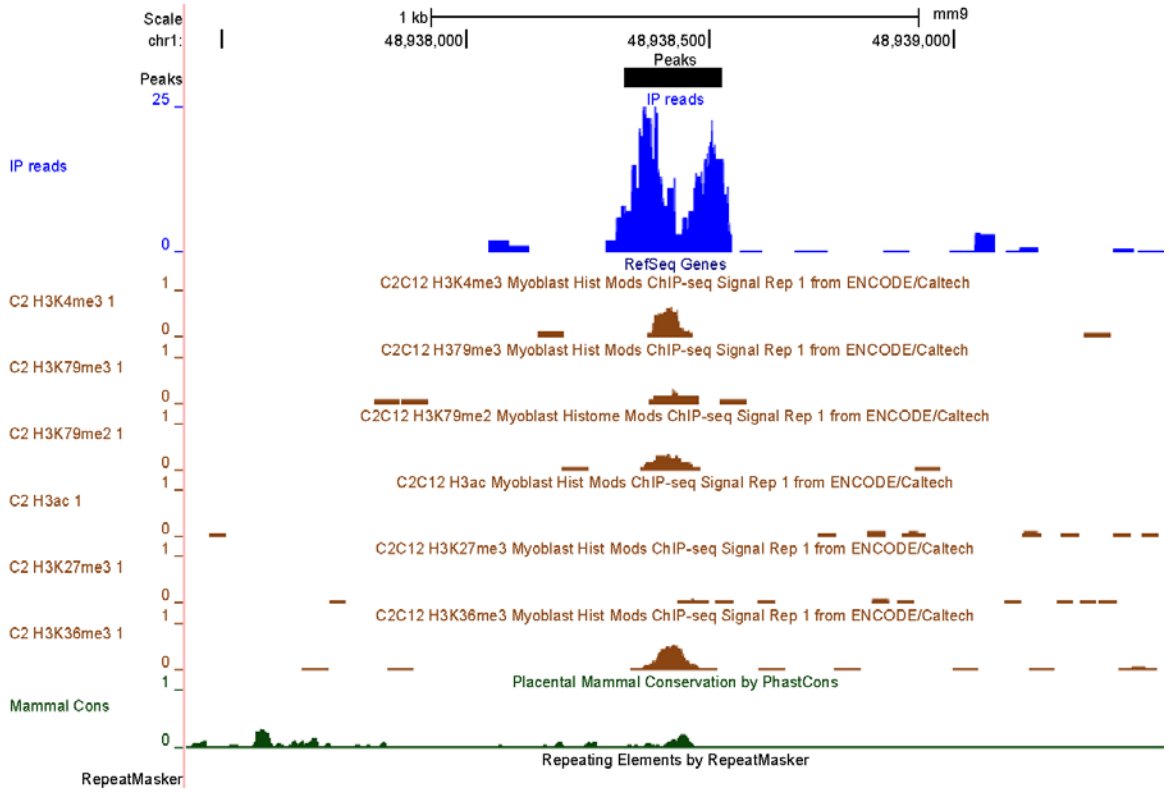


Figure 3.3 A PGC-1α peak overlapping the histone marks sequenced by the ENCODE Consortium.

3.3. *De novo* motif finding

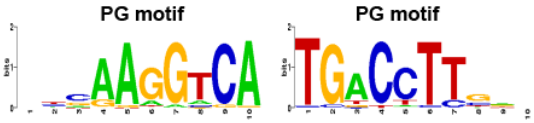
In Computational Biology, a sequence motif is a mathematical representation of the DNA-binding specificity of proteins such as nucleases and transcription factors [8]. Briefly, a motif can be built by aligning the known binding sites for a given factor and counting the frequencies of each nucleotide at each position; this matrix of frequencies is then converted in a matrix of probability scores, the Position Weight Matrix (PWM), whose logarithmic values are proportional to the binding energy contribution of the single bases [9] and can be represented by a sequence logo [10].

Motif analysis of the genomic regions identified by peak calling is a multi-purpose step in ChIP-Seq studies. In fact, when the motif of the chipped protein is already known, motif search can be used to validate the success of the experiment and to create a refined version of the motif, using the top scoring peaks to build a new one. On the other hand, when a motif of the chipped protein is not known, this analysis step can be useful for the determination *de novo* of the protein's binding specificity; likewise, in case of a co-regulator that does not bind DNA, motif analysis can provide an extremely valuable hint to: (i) discover new binding partners for the protein of interest, (ii) to verify the presence of binding sites for transcription factors that are known to work in complex with the chipped protein, and (iii) to identify which of the known partners are co-binding at each peak.

In the context of our ChIP-Seq experiments, since PGC-1α does not bind DNA directly, the goal of the motif analysis was rather focused on the identification of known and novel binding partners for the chipped coactivator. To this purpose, we used a *de novo* motif finder, PhyloGibbs [11], which combines a Gibbs motif sampler approach with phylogenetic footprinting into one integrated Bayesian framework. More in detail, the algorithm takes into account the phylogenetic relationships in an input set of aligned sequences to assign posterior probabilities to all possible configurations of an arbitrary number of binding sites for a user-specified number of unknown transcription factors.

When we used PhyloGibbs on the top scoring PGC-1α peaks, we identified a motif strongly resembling that of ERRα as the most enriched motif in the input dataset (for more details, see chapter 4). This result was confirmed when PhyloGibbs was run on the top ERRα peaks using different setting configurations; the combination of parameters z=1 y=200 m=10 – with "z" indicating the expected number of different motifs, "y" the number of expected sites and "m" the length of the expected motif – was the one that produced the motif with the most significant p-

value, 6.09726e-28. This motif was very similar to the one coming from the PGC-1α peaks and was recognized as highly similar to ESRRA.p2 (the weight matrix of ERRα) by the motif comparison software STAMP [12], as shown in Figure 3.4.



forward *reverse compliment*

Name	*E value*	*Alignment*	*Motif*
ESRRA.p2	7.7834e-10	TGACCTTNNN TGACCTTSAN	ESRRA.p2

Figure 3.4 The motif (top sequence logos) produced by PhyloGibbs using the top 200 ERRα peaks was recognized by the software STAMP as most similar to the SwissRegulon weight matrix ESRRA.p2 (bottom right sequence logo).

Although we expected to find this motif as the top result of the *de novo* search in ERRα peaks, we were surprised to observe it emerging uncontested from the PGC-1α peaks as well; in fact, although for ERRα the most notable partner is PGC-1α, this coactivator is known to interact with numerous other nuclear receptors, including PPARG, GR, NRFs, TFAM and RXRs (for a review on PGC-1α partners, see [13]), which nevertheless were not as dominant as ERRα. However, it needs to be considered that, since the high similarity between nuclear receptor binding preferences is also reflected in their motifs, it is likely that at least some fraction of the ESRRA binding sites were actually belonging to other similar nuclear receptors.

3.4. Binding sites prediction and motif over-representation compared to a shuffled background

In addition to the *de novo* motif search, we also performed binding site prediction for known transcription factors using the software MotEvo [14]. Briefly, by taking as input a multiple alignment of orthologous DNA sequences from different species, a collection of PWMs and a phylogenetic tree relating the species to the reference, MotEvo estimates the posterior probability for a given transcription factor binding site (TFBS) to occur at each position within the input sequences. The main advantage of this software consists in improving the accuracy of

binding site predictions by incorporating conservation information and by introducing a novel model, named “unknown functional element” (UFE), that takes into account the probability for an unknown factor for which we do not yet have a motif, to occur and compete with any of the input motifs.

As described more in detail in chapter 4, we aligned the peak regions to six other mammalian species and run MotEvo on this multiple alignment using a collection of 190 regulatory motifs downloaded from the SwissRegulon portal [2]. In parallel, we built a background set of random peaks to assess the significance of our predictions and to identify the over-represented motifs in our dataset. Instead of using a random set of genomic regions, which might not have had the same conservation pattern of our peaks, we shuffled the columns of the peak multi-alignments, in order to maintain both the gap and the conservation pattern of the original sequences. Moreover, because of the shuffle, the output regions were also size- and number-matched to the input peaks. As for the original peaks, we ran MotEvo on the shuffled peaks and we kept only those binding sites with a posterior probability greater than 0.1.

For each motif x, we calculated the sum of its binding site posterior probabilities in the peak alignments and in the shuffled dataset as:

$$n_x = \sum_i P(x_i)$$

and:

$$n'_x = \sum_j P(x_j)$$

where i and j are the i-th and j-th binding sites found in the original peaks and in the shuffled peaks, respectively. For both sets of peaks, we calculated the quantities f_x and f'_x, corresponding to the fraction of basepairs occupied by the motif x and weighted by its TFBSs posterior sum, as:

$$f_x = \frac{n_x l_x}{L_x}$$

and:

$$f'_x = \frac{n'_x l_x}{L'_x}$$

where l_x is the length of motif x, while L_x and L'_x are the total lengths of the original and shuffled peaks. Finally, the enrichment of TFBSs for each motif x was computed as:

$$Z = \frac{f_x - f'_x}{\sqrt{\frac{f_x * (1 - f_x)}{L_x} + \frac{f'_x * (1 - f'_x)}{L'_x}}}$$

Interestingly, when we checked which motifs were the most enriched in three subsets of PGC-1α peaks assigned to up-regulated, non-changing and down-regulated promoters, ESRRA turned out to be the most enriched motif for every subset (see Table 3.2), suggesting a key role of the ERRα nuclear receptor in the PGC-1α-controlled regulatory network. In addition, it is worth to notice that the group of leucine zipper transcription factors including FOS, JUN, BACH2, NFE2, NFE2L2 and FOSL2 were among the top over-represented motifs only in peaks assigned to up-regulated and to non-changing promoters, indicating that they might not be involved in the repression of PGC-1α downstream targets. Two transcription factors enriched in PGC-1α peaks assigned to down-regulated promoters, FOXJ2 and TFAP4, were previously found to play a role in preventing and contrasting the deleterious effects of inflammation [15, 16]. However, since the Z-scores are much smaller than in the other two peak sets, we believe that these peaks might be "spurious" and, thus, that the associated promoters are not direct targets of PGC-1α.

Table 3.2 Top over-represented motifs in three subsets of the PGC-1α peaks. Binding site predictions were compared to those predicted in a background set built by shuffling the original peak alignments.

PGC-1α peaks assigned to up-regulated promoters		**PGC-1α peaks assigned to non-changing promoters**		**PGC-1α peaks assigned to down-regulated promoters**	
Motif name	**Z-score**	**Motif name**	**Z-score**	**Motif name**	**Z-score**
ESRRA	30.81	ESRRA	46.71	ESRRA	8.93
NR5A1,2	26.20	NR5A1,2	40.87	NR6A1	5.92
NR6A1	20.43	BACH2	38.78	FOX{I1,J2}	5.81
BACH2	16.48	NR6A1	35.64	NFIX	5.78
FOS_FOS{B,L1}_JUN{B,D}	15.36	FOSL2	34.15	TFAP4	5.77
FOSL2	14.87	FOS_FOS{B,L1}_JUN{B,D}	33.31	NR5A1,2	5.73
SP1	13.18	NFE2	29.15	CDX1,2,4	5.12
NFE2L2	12.67	CTCF	26.30	MYFfamily	4.72
NFE2	12.67	HNF4A_NR2F1,2	25.20	BPTF	4.71
CTCF	12.54	NFE2L2	25.18	NFE2	4.59
HNF4A_NR2F1,2	12.44	REST	21.31	TLX1..3_NFIC{dimer}	4.07
NR1H4	11.84	NR1H4	19.22	ETS1,2	4.04
NFY{A,B,C}	11.65	PPARG	18.75	PPARG	4.03
PPARG	11.29	MZF1	18.65	TFCP2	3.99
MAZ	11.18	RREB1	17.89	PATZ1	3.90
ELF1,2,4	11.14	TFAP4	17.19	ATF2	3.69
EBF1	10.88	PATZ1	16.87	AIRE	3.59
PATZ1	10.65	ESR1	16.60	MZF1	3.56
TFCP2	10.37	SPI1	16.36	NFE2L2	3.48
RXRG_dimer	10.33	RXRA_VDR{dimer}	15.86	PDX1	3.43

3.5. Principal Component Analysis on the peaks site count matrix

Usually, the outcome of motif analysis provides an overview of all the motifs for which at least one binding site was predicted within the input regions of interest. In the case of ChIP-Seq data, considering that peaks are in general much wider than the actual protein binding site, this translates in the possibility of finding a TFBS for almost any known sequence motif. However, even by defining a score for motif over-representation versus a background set of regions, it is not trivial where to set the cutoff for the identification of the critical group of transcription factors responsible for the major effects observed in a dataset. Moreover, it has to be considered that the results of the motif enrichment analysis will be strongly influenced by the chosen background dataset.

For these reasons, we decided to use a background-independent approach, Principal Component Analysis (PCA), to grasp the most relevant information contained in our multi-dimensional dataset, at nearly no computational cost. Although PCA has been previously used for gene selection in microarrays and to compare two biological conditions in ChIP-Seq data [17, 18], to our knowledge it has not been employed so far to identify the structure of binding site occurrences in ChIP-Seq peaks. To this purpose, we built a "*site count matrix*" X containing the total number of predicted binding sites X_{pm} for each of the 190 mammalian weight matrices (columns) downloaded from SwissRegulon [2] in each peak p (rows). For each motif, the average site count was subtracted from every value, in order to mean-center the columns of the matrix:

$$\tilde{X}_{pm} = X_{pm} - \langle X_m \rangle$$

Then, a Singular Value Decomposition (SVD) was applied to the $\tilde{X}_{pm}$ matrix, to factorize it as:

$$\tilde{X}_{pm} = U \cdot S \cdot V^T$$

where U is a $m \times n$ matrix whose columns are the left singular vectors (in our case, the "*eigenpeaks*") of $\tilde{X}_{pm}$; S is a $n \times n$ diagonal matrix containing the singular values, and V^T (the transpose of V) is an $n \times n$ matrix whose rows are the right singular vectors (in our case, the "*eigenmotifs*"), with m the number of peaks, and n the number of motifs (Fig.3.5). The singular vectors of singular values are called principal *components*, which are linear combinations of uncorrelated variables and correspond to the number of original variables. Usually, only a few

components, accounting for most of the variation observed in the original dataset, are retained for further interpretation of the results.

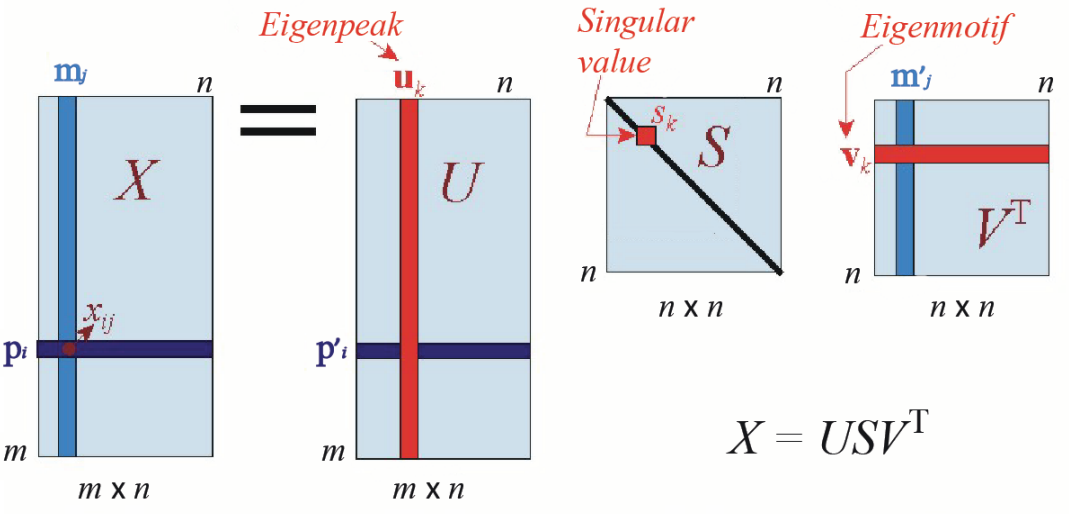


Figure 3.5 Schematic representation of the singular value decomposition. *Adapted from [19]*.

In our context, before applying the PCA on PGC-1α and ERRα peaks (see chapters 4 and 5), we tested its effectiveness on a group of ChIP-Seq experiments for several well studied transcription factors from the ENCODE Project, in order to confirm the results with the known interacting partners from the literature. To this end, we downloaded 20 peak sets in the GM12878 human lymphoblastoid cell line from the Crunch webserver (http://crunch.unibas.ch/ENCODE_REPORTS/), a recently developed and completely automated tool for the analysis of ChIP-Seq data which implements the read quality control algorithm described in chapter 2 and performs read mapping, peak detection and *de novo* motif analysis. After running MotEvo [14] on the peaks, we created a site count matrix as previously described and used R [20] to perform the PCA through SVD. In most of the cases, we could identify some already known partners of the chipped protein among the outliers of the PCA analysis. For example, Figure 3.6 shows the result of the PCA for the peaks of c-MYC, a member of the basic Helix-Loop-Helix (bHLH) family of transcription factors; as illustrated by the plot, the bHLH_family motif and the ARNT_ARNT2_BHLHB2_MAX_MYC_USF1 motif (which includes c-MYC, but also its well-known partner MAX) are pointing out of the motif cloud as the top relevant ones, followed by the motifs of other known c-MYC partner proteins, such as SP1 [21], MAZ [22], SREBP1 and 2 [23], HIF1α [24]. In conclusion, these results suggest that the PCA can be successfully applied to identify transcription factors which can be chosen for follow-up experiments as putative partners for the protein of interest.

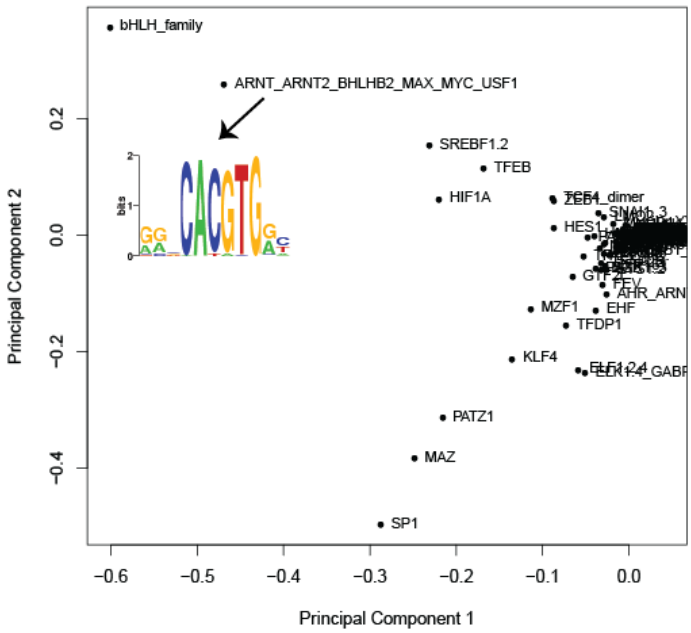


Figure 3.6 Principal Component Analysis of the peaks analyzed using Crunch on the c-Myc ENCODE ChIP-Seq data. The sequence logo of c-MYC is indicated by the arrow.

3.6. Evaluation of motif pairs preferred distances and geometries

Several recent studies have shown that pairs of interacting factors often co-occur along the genome with precise spatial preferences, which can be used as a criterion to predict *cis*-regulatory elements [25-27]. In order to determine whether certain transcription factors have a recurring or preferred distance from the chipped protein, we collected the distances within our peaks between binding sites for different motif pairs and calculated the entropy of the resulting distribution. As the entropy is a measure of the uncertainty of a variable, in general, the less uniform is a distribution, the lower will be its entropy; in our context, we expect to obtain a low entropy for the case a given transcription factor will occur at a preferred distance, rather than at random intervals.

In particular, since for some motifs we observed too few pair distances to use the classical Shannon entropy formulation, we adopted a Bayesian approach to estimate the predictability of the observed distances distribution. More in detail, to circumvent the case of having zero occurrences for certain distances (which would be incompatible with Shannon entropy), we need to introduce a pseudo-count.

Considering the discrete variable d of distances between motif pairs, the likelihood $P(d|\vec{\rho})$ of observing a certain distance d given a vector of probabilities $\vec{\rho}$, can be modelled as a multinomial distribution:

$$P(d = Data|\vec{\rho}) = \prod_{i=1}^{M} {\rho_i}^{n_i}$$

where ${\rho_i}^{n_i}$ is the probability of observing n_i times the distance $d = i$, with i ranging from 1 to the maximum distance M. The likelihood of the distances independent on $\vec{\rho}$ is called "*evidence"* and is calculated as:

$$P(d) = \int P(d|\vec{\rho})\, P(\vec{\rho})\, d\vec{\rho}$$

where $P(\rho)$ is the prior probability, that reflects our initial belief before having observed any data. We modelled the prior by a Dirichlet distribution:

$$P(\vec{\rho}|\alpha) \approx Dir(\alpha) = \frac{1}{\mathrm{B}(\alpha)} \prod_{i=1}^{M} {\rho_i}^{\alpha_\mathrm{i}-1}$$

where $\mathrm{B}(\alpha)$ is the beta function, which can be expressed in terms of the gamma function:

$$\mathrm{B}(\alpha) = \frac{\prod_{i=1}^{M} \Gamma(\alpha_i)}{\Gamma(\sum_{i=1}^{M} \alpha_\mathrm{i})}, \qquad \alpha = (\alpha_i, \dots, \alpha_M)$$

and is dependent on the hyperparameter α, which can be interpreted as a pseudo-count representing the expected number of distance counts a priori. When $\alpha = 1$ for all i, the prior corresponds to a uniform distribution.

Combining $P(d|\vec{\rho})$ and $P(\vec{\rho})$, we have:

$$P(d) = \int \Gamma(M\alpha) \prod_{i} \left[\frac{(\rho_i)^{n_i+\alpha-1}}{\Gamma(\alpha)}\right] d\vec{\rho}$$

which can be solved analytically as:

$$P(d) = \frac{\Gamma(M\alpha)}{\Gamma(n+M\alpha)} \prod_{i} \left[\frac{\Gamma(n_i+\alpha)}{\Gamma(\alpha)}\right]$$

If we take the logarithm of the above formula, it becomes:

$$log[P(d)] = log[\Gamma(M\alpha)] - log[\Gamma(n + M\alpha)] - Mlog[\Gamma(\alpha)] + \sum_i log[\Gamma(n_i + \alpha)]$$

Considering the uniform parameter prior ($\alpha = 1$) and the Stirlings' approximation:

$$log[\Gamma(x + 1)] \approx xlog(x) - x$$

then, by replacing ρ_i with (n_i/n), we can approximate the logarithm of the evidence as the negative entropy of the maximum likelihood (ML) estimate [28]:

$$log[P(d)] \approx -nH(n_i/n)$$

which can be derived such that the entropy H can be calculated as:

$$H = -\left[log[\Gamma(M\alpha)] - log[\Gamma(n + M\alpha)] - Mlog[\Gamma(\alpha)] + \sum_i log[\Gamma(n_i + \alpha)]\right]\frac{1}{n}$$

The value of H will be higher for more disordered distributions and lower for distributions showing a preferred distance. Indeed, as shown in Figure 3.7, we observed that motifs with very low entropy, like NF-κB, show a "spiked" distribution of distances from a reference motif (in this case, ESRRA), while motifs more ubiquitous like YY1 show a wider distribution, with higher entropy.

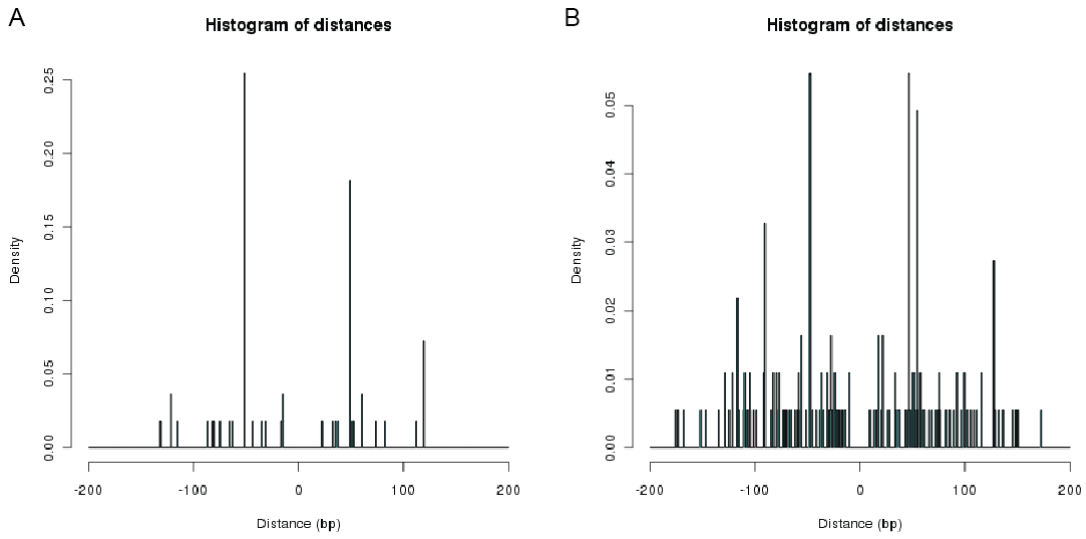


Figure 3.7 Examples of distributions of the distances between ESRRA and (**A**) NF-κB (H=5.23) or (**B**) YY1 (H=5.86).

However, even using this estimation of the entropy does not avoid labeling some distributions as less entropic than others simply because of a lower number of predicted binding sites, which favors random spikes with respect to real preferred distances. A possible solution to this problem, which we will work on in the next steps of this study, would be smoothing these counts by distributing the number of observed pair distances to their neighboring positions, either using an exponential kernel or by a simple discrete approximation.

Another interesting analysis of motif pairs is the study of the preferred geometries between their coupled binding sites. Several studies have shown that TFs bound to the same side of the DNA helix are much more likely to interact with each other and are more stable than proteins bound on opposite sides (for example, see [29, 30]). Moreover, a study on the *Drosophila* transcription factors Bcd and Kr showed that the maximal correlation between the two motifs was observed when the two TFs were located on the same DNA strand and shifted by two bases [31]. Considering that a complete turn of the DNA helix is composed of 10 nucleotides [32] and that zinc finger proteins (e.g. FOXP3) show enhanced binding to DNA when their dimers are located at regular turns [33], we studied which TFs were more often located at regular 10 bp distances from ERRα sites, up to four turns of the DNA helix. We decided to not consider longer distances, as higher-order conformations of the DNA helix might arise, thereby altering this short-range regularity. In order to count only TFBS pairs on the same face of the DNA, we considered only sites located on the same DNA strand.

Table 3.3 shows the results obtained using the “only ERRα” peaks (i.e. ERRα sites not directly co-activated by PGC-1α). Several motifs (KLF4, ZNF148, NR1H4, KLF12, FOX(C1,C2), SP1) are characterized by zinc finger domains, like ESRRA. KLF4 in particular, might occupy the same sites of SP1, as it was shown to compete with SP1 for binding to the DNA [34]. Interestingly, we noticed that most of the factors were more often found at a distance of 1 or 2 turns rather than at further distances, suggesting that these intervals reflect the appropriate distance for protein dimerization. Moreover, the fact that several TFs (for instance NR1H4 in turn 1) showed distinct preferences regarding binding before or after ESRRA sites might suggest a preferred orientation of their binding sites in concomitance with ERRα ones; these different preferences were present also when considering the group of zinc finger, despite their common domain type. In conclusion, we started investigating the patterns in which couples of binding sites are distributed; in the future developments of this project, we could explore these patterns more in depth and on a bigger set of sequences, in order to get more insights into transcription factor pair co-recruitment to DNA.

Table 3.3 TFBS occurrences at regular DNA turns relative to the center of the ESRRA binding sites in "only ERRα" peaks. "A" and "B" indicate whether the TFBS of the motif is after or before the TFBS of ESRRA, respectively. One turn corresponds to a distance of 10 bp, with a tolerance of ±2 bp. Counts are sorted by the column "turns 1+2", which should reflect the appropriate distance for dimerization.

MOTIF	1 turn B	1 turn A	2 turn B	2 turn A	3 turn B	3 turn A	4 turn B	4 turn A	turns 1+2	turns 1+2+3	turns 1+2+3+4
KLF4	5	1	11	36	3	4	7	2	53	60	69
ZNF148	14	30	4	3	3	2	3	2	51	56	61
NR1H4	28	7	4	6	7	5	6	3	45	57	66
KLF12	6	29	2	6	5	7	6	6	43	55	67
FEV	5	7	17	6	6	4	1	1	35	45	47
FOX{C1,C2}	2	28	2	2	2	5	1	0	34	41	42
NFE2L1	24	2	2	5	3	2	3	5	33	38	46
SP1	5	2	13	12	3	3	6	1	32	38	45
TP53	0	29	3	0	1	0	0	1	32	33	34
SPIB	8	4	17	2	3	6	13	7	31	40	60

References

1. Pruitt, K.D., et al., *RefSeq: an update on mammalian reference sequences.* Nucleic acids research, 2014. 42(Database issue): p. D756-63.
2. Pachkov, M., et al., *SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates.* Nucleic acids research, 2013. 41(Database issue): p. D214-20.
3. Bernstein, B.E., et al., *An integrated encyclopedia of DNA elements in the human genome.* Nature, 2012. 489(7414): p. 57-74.
4. Stadler, M.B., et al., *DNA-binding factors shape the mouse methylome at distal regulatory regions.* Nature, 2011. 480(7378): p. 490-5.
5. Mercer, T.R., et al., *DNase I-hypersensitive exons colocalize with promoters and distal regulatory elements.* Nature genetics, 2013. 45(8): p. 852-9.
6. Heintzman, N.D., et al., *Distinct and predictive chromatin signatures of transcriptional promoters and enhancers in the human genome.* Nature genetics, 2007. 39(3): p. 311-8.
7. Ong, C.T. and V.G. Corces, *Enhancer function: new insights into the regulation of tissue-specific gene expression.* Nature reviews. Genetics, 2011. 12(4): p. 283-93.
8. D'Haeseleer, P., *What are DNA sequence motifs?* Nature biotechnology, 2006. 24(4): p. 423-5.
9. Berg, O.G. and P.H. von Hippel, *Selection of DNA binding sites by regulatory proteins. Statistical-mechanical theory and application to operators and promoters.* Journal of molecular biology, 1987. 193(4): p. 723-50.
10. Schneider, T.D. and R.M. Stephens, *Sequence logos: a new way to display consensus sequences.* Nucleic acids research, 1990. 18(20): p. 6097-100.
11. Siddharthan, R., E.D. Siggia, and E. van Nimwegen, *PhyloGibbs: a Gibbs sampling motif finder that incorporates phylogeny.* PLoS computational biology, 2005. 1(7): p. e67.
12. Mahony, S. and P.V. Benos, *STAMP: a web tool for exploring DNA-binding motif similarities.* Nucleic acids research, 2007. 35(Web Server issue): p. W253-8.
13. Ventura-Clapier, R., A. Garnier, and V. Veksler, *Transcriptional control of mitochondrial biogenesis: the central role of PGC-1alpha.* Cardiovascular research, 2008. 79(2): p. 208-17.
14. Arnold, P., et al., *MotEvo: integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences.* Bioinformatics, 2012. 28(4): p. 487-94.
15. Chen, X., et al., *FOXJ2 expression in rat spinal cord after injury and its role in inflammation.* Journal of molecular neuroscience : MN, 2012. 47(1): p. 158-65.
16. Andriamanalijaona, R., et al., *Mediation of interleukin-1beta-induced transforming growth factor beta1 expression by activator protein 4 transcription factor in primary cultures of bovine articular chondrocytes: possible cooperation with activator protein 1.* Arthritis and rheumatism, 2003. 48(6): p. 1569-81.
17. Wang, A. and E.A. Gehan, *Gene selection for microarray data analysis using principal component analysis.* Statistics in medicine, 2005. 24(13): p. 2069-87.
18. Ji, H., et al., *Differential principal component analysis of ChIP-seq.* Proceedings of the National Academy of Sciences of the United States of America, 2013. 110(17): p. 6789-94.
19. Berrar, D.P., W. Dubitzky, and M. Granzow, *A practical approach to microarray data analysis.* 2003, Boston, MA: Kluwer Academic Publishers. xv, 368 p.
20. R Development Core Team, *R: A language and environment for statistical computing*, 2011, R Foundation for Statistical Computing: Vienna, Austria.
21. Parisi, F., P. Wirapati, and F. Naef, *Identifying synergistic regulation involving c-Myc and sp1 in human tissues.* Nucleic acids research, 2007. 35(4): p. 1098-107.

22. Bossone, S.A., et al., *MAZ, a zinc finger protein, binds to c-MYC and C2 gene sequences regulating transcriptional initiation and termination.* Proceedings of the National Academy of Sciences of the United States of America, 1992. 89(16): p. 7452-6.
23. Dang, C.V., *MYC, metabolism, cell growth, and tumorigenesis.* Cold Spring Harbor perspectives in medicine, 2013. 3(8).
24. Yu, Y., et al., *Mitochondrial regulation by c-Myc and hypoxia-inducible factor-1 alpha controls sensitivity to econazole.* Molecular cancer therapeutics, 2008. 7(3): p. 483-91.
25. Yokoyama, K.D., U. Ohler, and G.A. Wray, *Measuring spatial preferences at fine-scale resolution identifies known and novel cis-regulatory element candidates and functional motif-pair relationships.* Nucleic acids research, 2009. 37(13): p. e92.
26. Kulakovskiy, I.V., et al., *A deeper look into transcription regulatory code by preferred pair distance templates for transcription factor binding sites.* Bioinformatics, 2011. 27(19): p. 2621-4.
27. Vardhanabhuti, S., J. Wang, and S. Hannenhalli, *Position and distance specificity are important determinants of cis-regulatory motifs in addition to evolutionary conservation.* Nucleic acids research, 2007. 35(10): p. 3203-13.
28. Minka, T.P., *Bayesian inference, entropy, and the multinomial distribution.* Technical report, available at http://research.microsoft.com/en-us/um/people/minka/papers/minka-multinomial.pdf, 2003.
29. Yu, M., et al., *GA-binding protein-dependent transcription initiator elements. Effect of helical spacing between polyomavirus enhancer a factor 3(PEA3)/Ets-binding sites on initiator activity.* The Journal of biological chemistry, 1997. 272(46): p. 29060-7.
30. Falvo, J.V., et al., *A dimer-specific function of the transcription factor NFATp.* Proceedings of the National Academy of Sciences of the United States of America, 2008. 105(50): p. 19637-42.
31. Makeev, V.J., et al., *Distance preferences in the arrangement of binding motifs and hierarchical levels in organization of transcription regulatory information.* Nucleic acids research, 2003. 31(20): p. 6016-26.
32. Alberts, B., *Molecular biology of the cell.* 4th ed2002, New York: Garland Science. xxxiv, 1548 p.
33. Koh, K.P., M.S. Sundrud, and A. Rao, *Domain requirements and sequence specificity of DNA binding for the forkhead transcription factor FOXP3.* PloS one, 2009. 4(12): p. e8109.
34. Kanai, M., et al., *Loss of Kruppel-like factor 4 expression contributes to Sp1 overexpression and human gastric cancer development and progression.* Clinical cancer research : an official journal of the American Association for Cancer Research, 2006. 12(21): p. 6395-402.

4. Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the regulation of the hypoxic gene program

(manuscript published in Molecular and Cellular Biology, August 2014, vol.34 no.16, 2996-3012)

Mario Baresic[1,#], Silvia Salatino[1,2,3,#], Barbara Kupr[1], Erik van Nimwegen[2,3,*], and Christoph Handschin[1,*]

[1]Focal Area Growth & Development and [2]Focal Area Computational & Systems Biology, Biozentrum, University of Basel, Basel 4056, Switzerland

[3]Swiss Institute of Bioinformatics, Basel 4056, Switzerland

*Correspondence to: Christoph Handschin, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: christoph.handschin@unibas.ch +41 (0)61 267 23 78 or Erik van Nimwegen, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: erik.vannimwegen@unibas.ch +41 (0) 61 267 15 76.

[#]These authors contributed equally to this manuscript

Short title: Transcriptional network regulation by PGC-1α

Keywords: skeletal muscle; PGC-1α; coregulators; transcriptional regulation; transcription factor binding site prediction; transcriptional networks; motif activity response analysis; principal component analysis; exercise

Declaration of contribution to this study:

- I contributed to the design and organization of the project;

- I performed all the computational analyses of the project, including: raw sequencing data quality filtering and mapping to the reference genome, identification of bound regions, peak annotation, expression array analysis, de novo motif finding and TFBSs over-representation, principal component analysis, gene ontology enrichment analysis, evaluation of motif activity at direct and indirect targets of PGC-1α;

- I contributed to the analysis and to the interpretation of the results;

- I contributed to the writing of the manuscript and to the creation of the figures;

- I contributed to the revision of the manuscript.

4.1. Abstract

Skeletal muscle tissue shows an extraordinary cellular plasticity, but the underlying molecular mechanisms are still poorly understood. Here we use a combination of experimental and computational approaches to unravel the complex transcriptional network of muscle cell plasticity centered on the peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α), a regulatory nexus in endurance training adaptation. By integrating data on genome-wide binding of PGC-1α and gene expression upon PGC-1α over-expression with comprehensive computational prediction of transcription factor binding sites (TFBSs), we uncover a hitherto underestimated number of transcription factor partners involved in mediating PGC-1α action. In particular, principal component analysis of TFBSs at PGC-1α binding regions predicts that, besides the well-known role of the estrogen-related receptor α (ERRα), the activator protein-1 complex (AP-1) plays a major role in regulating the PGC-1α-controlled gene program of hypoxia response. Our findings thus reveal the complex transcriptional network of muscle cell plasticity controlled by PGC-1α.

4.2. Introduction

A sedentary life style can lead to an imbalance between energy intake and expenditure and favors the development of a number of chronic diseases like obesity and type 2 diabetes. Regular exercise on the other hand is an effective way to reduce the risk for these lifestyle-related pathologies [1]. The health benefits of exercise are at least in part induced by changes in skeletal muscle tissue. Muscle cells exhibit a high plasticity and thus a remarkably complex adaptation to increased contractile activity. For example, endurance training induces mitochondrial biogenesis, increases capillary density and improves insulin sensitivity [1, 2]. To achieve such a complex plastic response, a number of different signaling pathways are activated in an exercising muscle, for example p38 MAPK-mediated protein phosphorylation events, increased intracellular calcium levels or the activation of the metabolic sensors AMP-dependent protein kinase (AMPK) and sirtuin-1 (SIRT1) [3].

While the temporal coordination of the numerous inputs is not clear, all of the major signaling pathways converge on the peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) to either induce Ppargc1a gene expression, promote post-translational modifications of the PGC-1α protein, or by doing both [4, 5]. Upon activation, PGC-1α mediates the muscular

adaptations to endurance exercise by coactivating various transcription factors (TFs) involved in the regulation of diverse biological programs such as mitochondrial biogenesis, angiogenesis, ROS detoxification or glucose uptake [3]. Accordingly, transgenic expression of PGC-1α in mouse skeletal muscle at physiological levels not only induces mitochondrial biogenesis but also drives a fiber type conversion towards a more oxidative, slow-twitch phenotype [6] while muscle-specific Ppargc1a knockout animals exhibit several symptoms of pathological inactivity [7, 8].

Coregulators are part of multicomponent regulatory protein complexes that are well suited to translate external stimuli into changes in promoter and enhancer activities by combining various enzymatic activities to modulate histones and chromatin structure, and recruit other TFs [9]. Thus, dynamic assembly of distinct coregulator complexes enables the integration of many different signaling pathways leading to a coordinated and specific regulation of entire biological programs by multiple TFs [10, 11]. For example, PGC-1α not only recruits histone acetylases [12], the TRAP/DRIP/Mediator [13] as well as the SWI/SNF protein complexes [14], but also binds to and coactivates a myriad of different transcription factors, even though a systematic inventory of TF binding partners has not been compiled yet [15]. Thus, the specific control exerted by the PGC-1α-dependent transcriptional network might provide an explanation for the dynamic and coordinated muscle adaptation to exercise. Since PGC-1α in skeletal muscle not only confers a trained phenotype, but also ameliorates several different muscle diseases [16], the unraveling of the PGC-1α-controlled transcriptional network in skeletal muscle would be of great interest to identify putative therapeutic targets within this pathway.

Therefore, we aimed at obtaining a global picture of the co-regulatory activity of PGC-1α in skeletal muscle cells. More precisely, by combining data on the genome-wide binding locations of PGC-1α and the gene expression profiles in response to PGC-1α over-expression with comprehensive computational prediction of transcription factor binding site (TFBS) occurrence, we sought to unveil the biological processes that are regulated by PGC-1α, to identify the transcription factors that partner with PGC-1α, and to determine the mechanistic details of PGC-1α-regulated transcription. We not only mapped the locations on the DNA where PGC-1α was bound, but also delineated the target genes whose expression is either directly or indirectly affected by PGC-1α and identified novel putative transcription factor partners that mediated PGC-1α's action. In particular, our results strongly suggest that the activator protein-1 (AP-1) complex is a major regulatory partner of PGC-1α, with AP-1 and PGC-1α together regulating the hypoxic response gene program in muscle cells in vitro and in vivo.

4.3. Materials and methods

4.3.1. Cell culture and siRNA transfection

C2C12 cells were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 Units/ml penicillin and 100ug/ml streptomycin. To obtain myotubes, the C2C12 myoblasts were allowed to reach 90% confluence and the medium was changed to DMEM supplemented with 2% horse serum (differentiation medium) for 72 hours.

The siRNAs for the knockdown of NFE2L2, FOS, JUN, ATF3, NFYC, ZFP143, GTF2I, the non-targeting siRNA pool and the DharmaFECT1 transfection reagent were purchased from Dharmacon (Fisher Scientific) and the siRNA transfection was performed according to the Thermo Scientific DharmaFECT Transfection Reagents siRNA Transfection Protocol. Briefly, after three days of differentiation, the respective siRNAs (50nM final concentration) was added to the medium. 24h after siRNA transfection, the cells were infected with either the PGC-1α or GFP adenovirus. Then, 48h after adenoviral infection, the cells were harvested.

Differentiated C2C12 cells were infected with adenoviral (AV) shERRα (kindly provided by Dr. Anastasia Kralli, Scripps Research Institute, La Jolla, CA, USA) to knockdown and inactivate ERRα or shGFP as a control. The infected cells were kept in culture for 4 days. Afterwards, cells were infected with the AV-flag-PGC-1α or AV-GFP and kept in culture two additional days. As a supplement to the previously infected AV shERRα cells, 2μM of the ERRα inverse agonist XCT-790 were added. To the remaining cells, 0.02% DMSO as a vehicle was added to the differentiated medium. All the experiments have been performed in biological triplicates. For RNA isolation, TRIzol® was used according to the TRIzol® reagent RNA isolation protocol (Invitrogen). Three conditions were used for further analysis: AV-shGFP + AV-GFP + vehicle, AV-shGFP + AV–flag-PGC-1α + vehicle, AV-shERRα + AV-flag-PGC-1α + 2μM XCT-790.

4.3.2. ChIP and ChIP Sequencing

ChIP was performed according to the Agilent Mammalian ChIP-on-chip Protocol version 10.0. For each immunoprecipitation, approximately $1x10^8$ C2C12 cells were differentiated into myotubes and infected with AV-flag-PGC-1α. For cross-linking protein complexes to DNA binding elements, the cells were incubated in a 1% formaldehyde solution for 10 minutes, followed by the addition of glycine to a final concentration of 125mM to quench the effect of the formaldehyde. The cells were rinsed in 1xPBS, harvested in ice-cold 1xPBS using a silicone scraper and pelleted by centrifugation. The pelleted cells were either used immediately or flash

frozen and stored for later. The cells were then lysed at 4°C using two lysis buffers containing 0.5% NP-40/0.25% Triton X-100 and 0.1% Na-deoxycholate/0.5% N-lauroylsarcosine, respectively. The chromatin was then sheared by sonication to obtain DNA fragments of about 100-600bp in length. 50µl of the sonicated lysate was saved as input DNA. The immunoprecipitation was performed overnight at 4°C using magnetic beads (Dynabeads® Protein G, Invitrogen), which were previously coated with monoclonal antibodies like the monoclonal ANTI-FLAG® M2 Antibody, Sigma for the ChIP of PGC-1α or with the monoclonal anti-c-Fos (9F6) rabbit antibody #2250, Cell Signaling for the ChIP of FOS. The beads carrying the precipitate were washed five times for the c-Fos antibody and six times for the flag antibody with RIPA buffer and once with TE that contained 50mM NaCl to eliminate unspecific binding of DNA to the beads. For elution, the beads were resuspended in elution buffer containing 1% SDS, placed in 65°C water bath for 15 minutes and vortexed every 2 minutes. To reverse the cross-links, the samples were incubated at 65°C overnight. The following day, the RNA and the cellular proteins were digested using RNase A and proteinase K. The DNA was precipitated and the success of the chromatin immunoprecipitation was validated by semiquantitative real-time PCR. The ChIP experiments were performed in triplicates. The ChIP of PGC-1α was further used for sequencing. The ChIP-Seq experiment of over-expressed PGC-1α in C2C12 cells was performed in biological duplicates. At the joint Quantitative Genomics core facility of the University of Basel and the Department of Biosystems Science and Engineering (D-BSSE) of the ETH Zurich in Basel, DNA libraries were prepared using the standard Illumina ChIP-Seq protocol, as described by the manufacturer, and the immunoprecipitated samples sequenced on the Genome Analyzer II. In order to keep only high quality data, the sequenced reads were filtered based on the quality score of each read and its alignments. Reads were retained when Phred score >= 20, read length >= 25 bps and number of wrongly called nucleotides (Ns) <= 2. Those reads that passed the filter, (6'711'717 for the first immunoprecipitated sample (IP), 36'580'431 for the second IP, 17'899'074 for the first Whole Cell Extract (WCE), and 35'525'221 for the second WCE), were aligned to the mouse genome, UCSC mm9 assembly, using Bowtie version 0.12.7 [17] using parameters --best --strata -a --m 100. The number of aligned reads equaled 5'699'648 for the first IP sample, 16'053'370 for the first WCE, 21'448'059 for the second IP, and 32'244'584 for the second WCE.

4.3.3. Identification of bound regions

To identify regions that were significantly enriched in the ChIP, we passed a 200 bps long sliding window along the genome, sliding by 25 bps between consecutive windows, and estimated the

fraction of all ChIP reads f_{IP} that fall within the window, as well as the fraction f_{WCE} of reads from the whole cell extract that fall in the same window (which we estimate from a 2000 bps long window centered on the same genomic location). A Z-score quantifying the enrichment in the ChIP of each window was computed as:

$$Z = \frac{f_{IP} - f_{WCE}}{\sqrt{\sigma^2{}_{IP} + \sigma^2{}_{WCE}}}$$

where $\sigma^2{}_{IP}$ and $\sigma^2{}_{WCE}$ are the variances of the IP and WCE read frequencies, which are given by:

$$\sigma^2{}_{IP} = \frac{f_{IP} * (1 - f_{IP})}{N_{IP}} \text{ and } \sigma^2{}_{WCE} = \frac{f_{WCE} * (1 - f_{WCE})}{N_{WCE}}$$

respectively.

The enrichments were reproducible across biological replicates. Using only the first sequencing dataset, we called peaks at a Z cutoff of 4.5; we then compared these with the Z scores from the corresponding regions of the second dataset and the Pearson correlation coefficient was found to be 0.778. Similarly, we called peaks at a Z cutoff of 4.5 using only the second sequencing dataset; when we compared these peaks with the Z scores of the corresponding regions from the first dataset, the Pearson correlation coefficient was found to be 0.782.

To obtain a final set of binding peaks, we combined the reads from the two biological replicates computing the Z score of each window was computed as:

$$Z = \frac{f_{IP1} + f_{IP2} - f_{WCE1} - f_{WCE2}}{\sqrt{\sigma^2{}_{IP1} + \sigma^2{}_{IP2} + \sigma^2{}_{WCE1} + \sigma^2{}_{WCE2}}}.$$

We conservatively considered all windows with a Z-score larger than 4.5 as were considered significantly enriched (False Discovery Rate 0.6%). The final binding peaks were obtained by merging consecutive windows that all passed the cut-off and by considering the “peak” to correspond to the top scoring window, i.e. corresponding to the summit of the ChIP-Seq signal. To determine the PGC-1α distribution genome-wide, peaks were annotated according in relation to their closest Mus musculus RefSeq transcripts. We defined peaks as: “Intronic” (peak center lying inside an intron); “Exonic” (peak center lying inside an exon); “Upstream of TSS” (peak center lying within -10 to 0 kb from the closest TSS); “Downstream of TES” (peak center lying within 0 to 10 kb from the closest TES); “Intergenic” (peak center located farer than 10 kb from the nearest transcript). Moreover, we computed the ratio between observed and expected peak

location distributions, obtained by generating 100 peak sets composed of 7512 random peaks each.

4.3.4. Motif finding and TFBSs over-representation

The binding peak regions were aligned to orthologous regions from other 6 mammalian species – human (hg18), rhesus macaque (rheMac2), dog (canFam2), horse (equCab1), cow (bosTau3) and opossum (monDom4) – using T-Coffee [18]. A collection of 190 mammalian regulatory motifs (position weight matrices or WMs) representing the binding specificities of approximate 350 mouse TFs (in many cases, sequence specificities of multiple closely-related TFs were represented with the same WM) were downloaded from the SwissRegulon website [19]. TFBSs for all known motifs were predicted using the MotEvo algorithm [20] on the alignments of all the 7512 peak sequences. Only binding sites with a posterior probability >= 0.1 were considered for the further steps of the analysis. In order to create a background set of regions to assess the overrepresentation of binding sites within our regions, we created randomized alignments by shuffling the multiple alignment columns, maintaining both the gap patterns and the conservation patterns of the original alignments. TFBSs were predicted on the shuffled alignments using the same MotEvo settings as for the original peak alignments. Over-representation of motifs in the PGC-1α binding peaks was calculated by comparing total predicted TFBS occurrence within binding peaks with the predicted TFBS occurrence in the shuffled alignments. We evaluated the enrichment of TFBSs for each motif *x* by collecting the sum n_x of the posterior probabilities of its predicted sites in the peak alignments as well as the corresponding sum n'_x in the shuffled alignments, and computed a Z-score:

$$Z = \frac{f_x - f'_x}{\sqrt{\frac{f_x * (1 - f_x)}{L_x} + \frac{f'_x * (1 - f'_x)}{L'_x}}}$$

where L_x and L'_x are the total lengths of the original and shuffled alignments, respectively, while f_x and f'_x are given by the equations:

$$n_x * l_x = f_x * L_x \text{ and } n'_x * l_x = f'_x * L'_x$$

with l_x the length of motif *x*.

4.3.5. Principal Component Analysis of TFBS occurrence in binding peaks

The input matrix *N* for the Principal Component Analysis (PCA) contained the total number of predicted binding sites N_{pm} in each of the 7512 binding peaks *p* (rows) for each of the 190 mammalian regulatory motifs *m* (columns). After mean centering the columns of this matrix, $\widetilde{N}_{pm} = N_{pm} - \langle N_m \rangle$, i.e. subtracting the average site count for each motif, Singular Value Decomposition (SVD) was used to factorize this matrix: $\widetilde{N} = U \cdot S \cdot V^T$, where U is a $P \times M$ matrix whose columns are the left singular vectors of $\widetilde{N}$; S is a $M \times M$ diagonal matrix containing the singular values, and V^T (the transpose of V) is an $M \times M$ matrix whose rows are the right singular vectors, with *P* the number of peaks, and *M* the number of motifs. The SVD was performed using the "svd" package of the "R" programming language.

4.3.6. Gene expression arrays

Whole-gene expression after 48 hours of transfection with adenovirus was measured in C2C12 cells with Affymetrix GeneChip® Mouse Gene 1.0 ST microarrays at the Life Science Training core facility of the University of Basel. Raw probe intensities were corrected for background and unspecific binding using the Bioconductor package "affy" [21]. Subsequently, probes were classified as expressed or non-expressed by using the "Mclust" R package [22] and, after removal of non-expressed probes, the intensity values were quantile normalized across all samples. Using mapping of the probes to the UCSC collection of mouse mRNAs, probes were then associated to a comprehensive collection of mouse promoters available from the SwissRegulon database [19]. The log2 expression level of a given promoter was calculated as the weighted average of the expression levels of all probes associated to it. Log2 expression levels were then compared between over-expressed PGC-1α and the control GFP sample; for each promoter, the change in expression level across the two conditions was measured by log2 fold change (log2FC), computed as the difference between the mean of the log2 values in PGC-1α and the mean of the log2 values in GFP. The significance of the expression change was assessed by a Z score, which was computed as:

$$Z = \frac{\bar{E}_{PGC1\alpha} - \bar{E}_{GFP}}{\sqrt{\frac{\sigma^2{}_{PGC1\alpha}}{n} + \frac{\sigma^2{}_{GFP}}{n}}}$$

where $n = 3$ was the number of replicate samples, $\bar{E}_{PGC1\alpha}$ is the mean log2 expression across the PGC-1α samples, $\bar{E}_{GFP}$ is the mean log2 expression across the GFP samples, and $\sigma^2{}_{PGC1\alpha}$ and $\sigma^2{}_{GFP}$ are the variances of log2 expression levels across the replicates for the PGC-1α and

control samples, respectively. Promoters were considered significantly up-regulated when log2FC >= 1 and Z >= 3, and significantly down-regulated when log2FC <= -1 and Z <= -3.

Peaks were assigned to promoters by proximity. To assign each peak to a promoter, we calculated the distance from the center of the peak to the center of neighboring promoters; whenever the peak was closer than 10 kb from at least one promoter, it was assigned to the nearest promoter and, thus, to its associated gene.

4.3.7. Gene Ontology enrichment analysis

Gene IDs were extracted from differentially regulated promoters and divided in four groups: up-regulated promoters with an assigned binding peak, up-regulated promoters without an assigned binding peak, down-regulated promoters with an assigned peak, and down-regulated promoters without an assigned peak. These four gene sets were used as input for the functional analysis tool FatiGO [23] to identify significantly over-represented Gene Ontology (GO) categories compared to all Mus musculus genes. Only GO terms having an FDR-adjusted p-value <= 0.05 were considered significant.

4.3.8. Motif activity at direct and indirect targets of PGC-1α

To integrate the information from the PGC-1α binding peaks, we extended MARA [24] to model the direct and indirect regulatory effects of PGC-1α. Given the input expression data and the computationally predicted binding sites, MARA infers, for each of 190 regulatory motifs *m*, the activity A_{ms} of the motif in each sample *s* when the motif occurs *outside* of a region of PGC-1α, and the activities A^*_{ms} of the motifs when they occur *within* a PGC-1α binding peak. That is, changes in the motif activities A_{ms} upon over-expression of PGC-1α indicate indirect regulatory effects of PGC-1α on each motif *m*, whereas changes in the motif activities A^*_{ms} reflect direct regulatory effects of PGC-1α as mediated by each motif *m*. For each promoter *p* that was not associated with any PGC-1α binding peak (which we denote indirect targets), we modeled its log-expression in sample *s,* e_{ps}, in terms of the predicted number of TFBSs N_{pm} that occur in the proximal promoter region (running from -500 to +500 relative to TSS) for each regulatory motif *m.* That is, MARA assumes the linear model:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N_{pm} A_{ms}$$

where c_p is the basal expression of promoter *p*, $\tilde{c}_s$ is a sample-dependent normalization constant, and A_{ms} is the regulatory activity of motif *m* in sample *s*, which is inferred by the

model. Formally, A_{ms} quantifies amount by which the expression of promoter *p* in sample *s* would be reduced if a binding site for motif *m* were to be deleted from the promoter.

For each "direct target" promoter *p* that has an associated PGC-1α binding peak, which we defined as promoters with a peak within 1 kb or with a peak within 100 kb that was highly conserved according to PhastCons score of the region [25], we model its expression in terms of the predicted TFBSs in the binding peak, i.e.:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N^*_{pm} A^*_{ms}$$

where N^*_{pm} is the number of predicted TFBSs for motif *m* in *the peak* associated with promoter *p*, and A^*_{ms} is the motif activity of regulator *m* in sample *s* when this motif occurs in the context of PGC-1α binding. That is, the inferred motif activities A_{ms} quantify the activities of regulatory motifs when they occur independent of PGC-1α binding, and the motif activities A^*_{ms} quantify the activities of motifs when they occur in a PGC-1α binding peak, i.e. the latter activities reflect direct effects of a PGC-1α while the former reflect indirect effects.

MARA predicts activities for 190 different mammalian regulatory motifs, associated with roughly 350 mouse TFs. Besides motif activities MARA also calculates error-bars δ_{ms} for each motif *m* in each sample *s*. Using these, MARA calculates, for each motif *m*, an overall significance measure for the variation in motif activities across the samples analogous to a z-statistic:

$$z_m = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\frac{A_{ms}}{\delta_{ms}}\right)^2}$$

For each motif we calculate both a z-score z_m associated with its indirect activity changes, and a z-score z^*_m associated with its direct activity changes. MARA also ranks the confidence on predicted target promoters of each motif by a Bayesian procedure that quantifies the contribution of that factor to explaining the promoter's expression variation by a Chi-squared value (for details, see [24]). The parameters used for motif stratification were: (i) the Z score z^*_m for direct activity changes, (ii) the Z score z_m for indirect motif activity changes, (iii) the Z score $\bar{z}^*_m$ for direct motif activity changes, computed by averaging the sample replicates and (iv) the Z score $\bar{z}_m$ for indirect motif activity changes, computed by averaging the sample replicates. The latter two measures were used to show which direction the motif activity changes when over-

expressing PGC-1α with respect to the control condition. All motifs *m* for which either the direct or indirect motif activities were changing significantly ($z \geq 2$) were subsequently selected.

4.3.9. De novo motif finding

PhyloGibbs [26] was used to identify *de novo* motifs across the 200 top enriched PGC-1α peaks. The parameters used were -D 1 -z 1 -y 200 -m 10, corresponding to searching on multiple alignments for a single motif of length 10 with a total of 200 sites. The resulting motif was scanned for similarity to the other known motifs from our dataset using STAMP [27], with settings: Pearson Correlation Coefficient for column comparison metric, Smith-Waterman for the alignment method, penalty of 0.5 and 0.25 for gap opening and gap extension, respectively.

4.3.10. Real-time PCR and target gene validation

Putative target genes of distinct transcription factor-PGC-1α combinations were chosen according to three criteria: first, positive transcriptional regulation by PGC-1α by more than 2 fold, second, presence of a PGC-1α binding peak within a 10 kb distance from the TSS and third, prediction of targeting by MARA with a positive Chi-squared score. The sequences of the primers used in real-time PCR experiments are depicted in Suppl. Table 1. Relative mRNA was quantified by qPCR on a StepOnePlus system (Applied Biosystems) using Power SYBR Green PCR Master Mix (Applied Biosystems).

The values are presented as the mean +/- SEM. A Student's t-test was performed and a p-value < 0.05 was considered as significant (*p<0.05, **p<0.01, ***p<0.001).

4.3.11. Animals

Mice were housed in a conventional facility with a 12-h night/12-h day cycle with free access to chow diet pellet and water. For the experiments, 22-23 week-old skeletal muscle-specific HSA-PGC-1α knockout (MKO) male mice and 8 week-old PGC-1α muscle-specific transgenic (TG) male mice were used as previously described [6-8]. All experiments were performed according to the criteria outlined for the care and use of laboratory animals and with approval of the veterinary office of the canton Basel and the Swiss authorities.

4.3.12. Treadmill running

Treadmill running was performed with the TG mice on the Columbus Instruments motorized treadmill with an electric shock grid. The mice were acclimatized to the treadmill and then let run

till exhaustion. The running protocol is as follows: 10m/min for 5min with an increase by 2m/min every 5min until 26m/min and an inclination of 5 degrees. The speed of 26m/min was kept until exhaustion of the mice [7, 28, 29]. Mice were killed and tissues were collected 3h after exercise.

4.3.13. RNA isolation of muscle tissue

Gastrocnemius and quadriceps were used to isolate RNA by TRIzol® according to the TRIzol® reagent RNA isolation protocol (Invitrogen).

4.4. Results

4.4.1. Broad recruitment of PGC-1α to the mouse genome

PGC-1α-dependent gene transcription has been studied in many different experimental contexts. In isolation, gene expression arrays however are unable to distinguish direct from indirect targets, or to reveal the genomic sites where PGC-1α is recruited to enhancer and promoter elements, i.e. by coactivating TFs that directly bind to the DNA. Thus, we first performed chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq) of PGC-1α in differentiated C2C12 mouse myotubes to identify the locations where PGC-1α is bound to the genome. To identify genomic regions that are significantly enriched in the ChIP, we slid a 200 bp window across the genome comparing the local ChIP read density with the read density from a background whole cell extract sample. We selected all regions with a Z-statistic larger than 4.5 as significantly enriched (FDR 0.6%, Suppl. Fig. S1A). Using this stringent cutoff, we identified 7512 binding regions for PGC-1α via interaction with a TF genome-wide, which include binding regions in the promoters of known PGC-1α target genes (Fig. 1A) such as medium-chain specific acyl-CoA dehydrogenase (*Acadm*) and cytochrome c (*Cycs*) [30, 31]. The enrichment of immunoprecipitated DNA fragments from the ChIP-Seq was validated for these and other PGC-1α target genes by semiquantitative real-time PCR (Fig. 1B). In absolute terms, the distribution of the ChIP-Seq peaks revealed that PGC-1α is mostly recruited at distal sites from the assigned targets and, to a lesser extent, to proximal regions of the gene or within an intronic sequence (Fig. 1C). However, when compared to randomly selected DNA regions of equal size and number, PGC-1α binding peaks occur twice as often within 10 kb upstream of the transcription start site (TSS).

In parallel to the ChIP-Seq experiment, we furthermore analyzed gene expression patterns in differentiated muscle cells both in control condition and under PGC-1α over-expression. Using a

reference set of mouse promoters [19] and associating microarray probes to promoters by mapping to known transcripts, we found 1566 promoters (corresponding to 984 genes) to be significantly up-regulated (log2 fold change >= 1; Z score >= 3) and 1165 promoters (corresponding to 727 genes) to be significantly down-regulated (log2 fold change <= -1; Z score <= -3). Thus, similar to previous reports, PGC-1α induced and repressed the transcription of almost the same number of genes, respectively, indicating that the physiological function of PGC-1α includes both the activation and inhibition of substantial numbers of genes.

To combine the DNA binding results from the ChIP-Seq with the data of the gene expression arrays, we then assigned ChIP-Seq peaks to the closest promoter (and the associated gene) within a maximum distance of 10 kb. In this way, about 30% of all peaks (2295 of 7512) could be associated with a target promoter. Inversely, for about 35% of all significantly up-regulated genes (341 of 984), a PGC-1α binding peak is found within 10 kb of the promoter. Since some of the up-regulated promoters may be regulated by more distal peaks, this is only a lower bound on the fraction of genes that are directly regulated. In stark contrast, only about 5% of all repressed genes harbor one or more PGC-1α DNA recruitment peaks in their vicinity (36 of 727) opposed to 95% indirectly down-regulated PGC-1α target genes (691 genes) (Fig. 1D). Moreover, the distribution of the distances between PGC-1α peaks and their associated promoters revealed a tight cluster of 532 peaks close to promoter regions for up-regulated, direct PGC-1α target genes (Fig. 1E) whereas the distribution of the 43 peaks associated to down-regulated genes was much wider, raising the possibility that the association of peaks to transcriptionally repressed genes might be spurious (Fig. 1F). In summary, the strong enrichment of binding peaks near up-regulated genes and the almost complete absence of binding peaks near down-regulated genes suggest that direct regulation of transcription by PGC-1α is almost exclusively activating. We note that there is a large fraction of binding peaks (75%) that are associated to target genes that do not significantly alter their expression. These peaks may have been wrongly assigned, their functionality may be dependent on additional factors not active in these cells, or they may simply be spurious binding events that are not functional. We next used this stratification of peaks and genes to study whether direct (i.e. with an associated binding peak) and indirect PGC-1α target genes exert different biological function and identified Gene Ontology (GO) terms that were over-represented in any of the four categories. First, we observed that the most significantly enriched functional categories for directly and indirectly up-regulated genes were those related to mitochondria, oxidative phosphorylation and energy production (Fig. 1G and Suppl. Fig. S1B). In contrast, GO analysis of indirectly down-regulated PGC-1α target genes revealed a high prevalence of terms related to inflammation and immune

response (Fig. 1H and Suppl. Fig. S1C). Assuming that the assignment of peaks to repressed genes is not spurious, the few directly repressed PGC-1α targets exhibit an enrichment in functions related to muscle contraction, in particular for genes that are linked to contractile and metabolic properties of glycolytic, fast-twitch muscle fibers (Fig. 1H and Suppl. Fig. S1D), as would be expected from the observed shift from glycolytic to oxidative fibers mediated by PGC-1α in muscle [6].

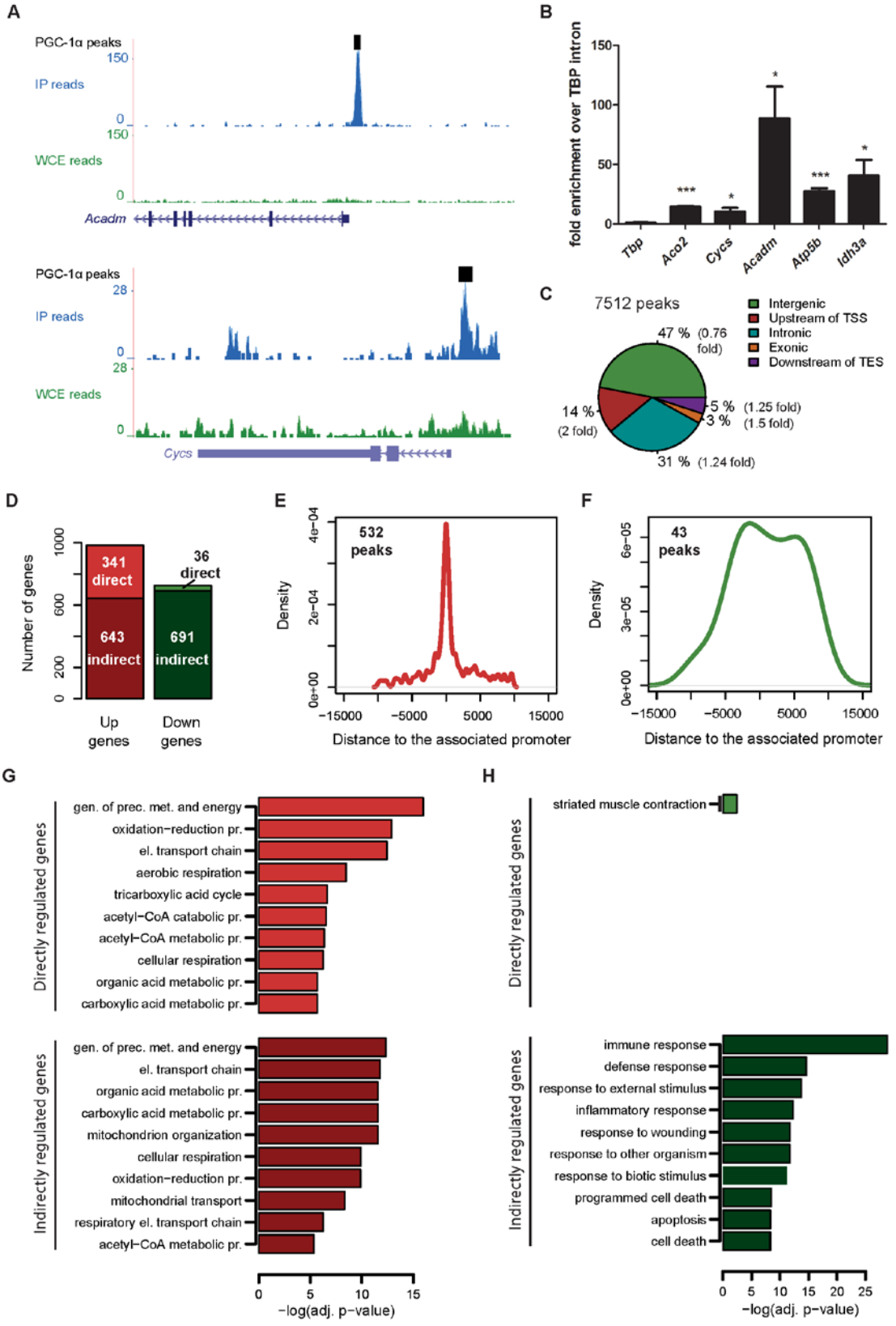


Figure 1. Genome-wide DNA recruitment of PGC-1α in mouse muscle cells. (**A**) PGC-1α ChIP-Seq binding peaks (read densities) around the TSS of the genes *Acadm* and *Cycs* obtained from the UCSC Genome Browser. (**B**) Real-time PCR validation of the ChIP enrichment measured at the promoter of a set

of PGC-1α target genes. Bars represent fold enrichment over that of the *Tbp* intron, error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001. (**C**) Mapping ChIP-Seq PGC-1α peaks across the genome. Transcription Start Site (TSS) and Transcription End Site (TES) are relative to mm9 RefSeq transcripts. "Intergenic": ≥ 10 kb from the nearest transcript; "Upstream of TSS": -10 to 0 kb from the TSS; "Downstream of TES": 0 to 10 kb from the TES. Numbers between brackets indicate, for each category, the ratio between the percentage of PGC-1α peaks and the percentage of the same number of randomly distributed peaks. (**D**) Histogram illustrating the number of direct and indirect genes either up- or down-regulated by over-expression of PGC-1α in muscle cells. Direct genes are those associated to promoters found within ±10 kb relative to the nearest peak. (**E**) Distribution of the distances of 532 peaks from their associated up-regulated gene promoters. (**F**) Distribution of the distances of 43 peaks from their associated down-regulated gene promoters. (**G-H**) Subset of the top significantly enriched GO Biological Process terms identified for directly and indirectly up-regulated (**G**) and down-regulated (**H**) PGC-1α target genes.

4.4.2. Modeling the direct and indirect gene regulatory effects of PGC-1α

As a next step, we rigorously modeled the effects of PGC-1α on its target genes in terms of the occurrence of TFBSs for a large collection of mammalian regulatory motifs. We previously introduced a general framework, called Motif Activity Response Analysis (MARA) [24], for modeling the gene expression profiles as a linear function of the TFBSs occurring in the promoters and unknown regulatory "activities" of each of the regulators. As detailed in the Methods, we here extended MARA to incorporate information from the PGC-1α ChIP-Seq data, with the aim of identifying which other TFs are involved in mediating both the direct and indirect regulatory effects of PGC-1α. Specifically, for all "direct target" promoters that were associated with a PGC-1α binding peak, we modeled the expression of the promoter in terms of the predicted TFBSs in the neighborhood of the binding peak, while for "indirect target" promoters we modeled the promoter's expression in terms of the predicted TFBSs in the proximal promoter region, according to the conventional MARA approach (Fig. 2A and 2B).

First, further supporting our analysis above, direct target promoters were almost exclusively up-regulated and only in a few exceptional cases reached statistical significance for PGC-1α-repressed transcripts (Fig. 2C). Among the direct motif activities, the ESRRA position weight matrix was the top ranking motif with a Z score of 6.04 (Suppl. Fig. 2). The corresponding TF estrogen-related receptor α (ERRα), an orphan nuclear receptor, has been extensively studied as a central binding partner for PGC-1α in the regulation of mitochondrial gene expression [30-33]. To stratify the different motifs according to their predicted function, we then divided all motifs into groups according to the behavior of both their direct and indirect activity changes. Strikingly, all motifs exhibited one of only four different motif activity patterns. First, 6 TFs (Suppl. Fig. S2) were predicted to positively regulate PGC-1α target genes only in the presence of PGC-1α (Fig. 2D). Second, we found 6 motifs (Suppl. Fig. S2) with significantly up-regulated direct and indirect motif activities upon PGC-1α over-expression (Fig. 2E). To our surprise, ERRα was predicted to

regulate PGC-1α target genes in this manner, even though in previous reports gene regulation by ERRα in the context of activated PGC-1α was suggested to be dependent on PGC-1α coactivation [30-33]. Third, we found 13 motifs (Suppl. Fig. S2) that were predicted to regulate PGC-1α target genes, however only in the absence of PGC-1α (Fig. 2F). Fourth, there was a group of 28 motifs (Suppl. Fig. S2) that showed a significant decrease of indirect motif activity upon PGC-1α over-expression, but no significant change of their direct motif activity, including NFκB (Fig. 2G), a central regulator of inflammation which is indirectly repressed by PGC-1α [33]. Intriguingly however, no motif was found that showed significant direct repression of target genes, reinforcing the hypothesis that PGC-1α-dependent gene repression is an indirect event.

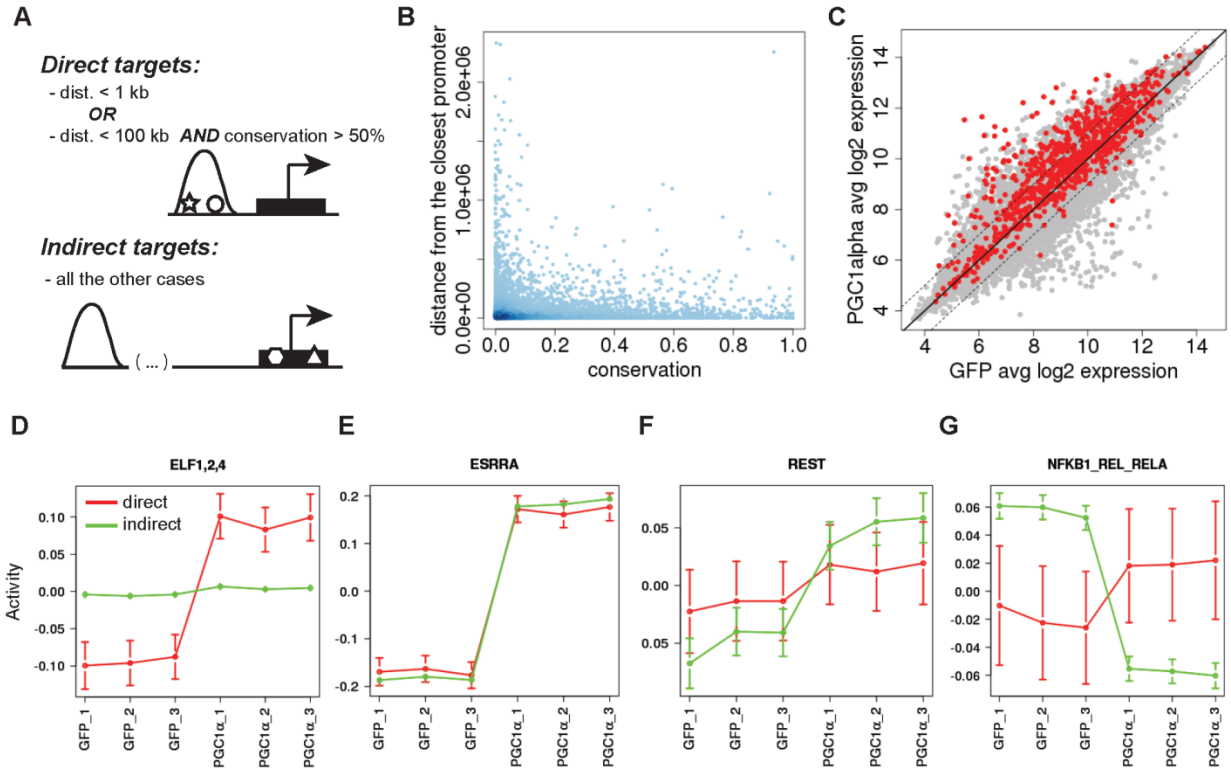


Figure 2. Four distinct mechanistic modes of action for gene expression regulated by PGC-1α and TF partners. (**A**) Classification of direct and indirect target genes in MARA (see Methods) (**B**) Distribution of peak distance from the closest promoter and phastCons conservation score of the peak. (**C**) Distribution of log2 expression values for all mouse promoters. Expression values were averaged across the 3 GFP and the 3 PGC-1α samples. Direct targets are depicted in red, indirect targets in grey. (**D-G**) Activity plot of the motifs ELF1,2,4 (**D**), ESRRA (**E**), REST (**F**) and NFKB1_REL_RELA (**G**) as predicted by MARA (Motif Activity Response Analysis). Red: direct targets; green: indirect targets.

4.4.3. Nuclear receptors and activator protein-1-like leucine zipper proteins are the main functional partners of PGC-1α in muscle cells

As a next step, we analyzed the occurrence of TF DNA-binding motifs in the PGC-1α peaks identified by ChIP-Seq. We first performed *de novo* motif prediction on the top 200 peaks, using PhyloGibbs [26]. As shown in Figure 3A, the motif that PhyloGibbs identified matches significantly (E-value = 7.7834e-10, as calculated by STAMP [27]) the canonical ESRRA motif. In addition to the *de novo* prediction, we also used the same collection of 190 mammalian regulatory motifs used by MARA [19] to check which known TF DNA-binding motifs were significantly over-represented in the PGC-1α peaks relative to a set of background regions.

Many of the most significantly enriched motifs represent variations of nuclear receptor binding sequences that are based on the "AG^{T}/$_{G}$TCA" core hexamer and occur either alone or in direct, inverted or everted repeats with variable spacing (Fig. 3B). Of these, the most significantly enriched motif was ESRRA, which is present in ~20% of all peaks. Moreover, among all genes with at least one associated binding peak within 10Kb, ~28% are associated with a peak containing a predicted ERRα site. Interestingly, besides the nuclear receptor motifs, we also found the DNA-binding element of the insulator protein CCCTC-binding factor (CTCF), and a set of highly similar DNA elements sharing the FOS-JUN-like recognition sequence "TGA^{G}/$_{C}$TCA" bound by the TFs BACH2, FOS, FOSB, FOSL1, JUN, JUNB, JUND, FOSL2, NFE2, and NFE2L2 among the top 15 motifs enriched in PGC-1α peaks (Fig. 3B).

The identity of the exact nuclear receptor binding partner that is bound at each peak is difficult to deduce from DNA-binding motifs, since considerable promiscuity exists between receptors and DNA-binding elements in different configurations of hexameric repeats [34]. Moreover, non-nuclear receptor-like TFs are less well studied in the context of PGC-1α-controlled gene expression. Thus, to identify which regulatory motifs are most over-represented among peaks that do not contain nuclear receptor-like sites, we first manually grouped all of the motifs with a sequence logo very similar to that of ESRRA. Next, we discarded all peaks that had one or more predicted TFBSs for any of the motifs in this set. With the remaining 3856 DNA sequences (51.33% of the peaks), we then again assessed the over-representation of each of the 190 mammalian regulatory motifs. In this analysis, "TGA^{G}/$_{C}$TCA" recognition elements, hence FOS-JUN-like motifs, were the most significantly enriched among these peaks (Fig. 3C). This result suggests that PGC-1α peaks naturally fall into two classes: those containing ESRRA-like sites, and those containing sites for FOS-JUN-like motifs.

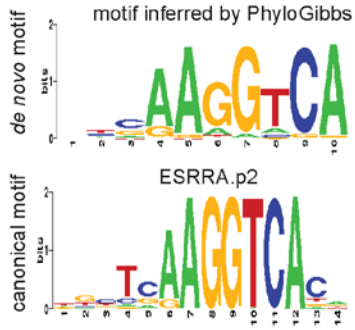
We then constructed a matrix *N*, whose elements N_{pm} contain the number of predicted TFBSs for each motif *m* in each peak region *p*. We then performed principal component analysis (PCA) on this site-count matrix to identify linear combinations of regulatory motifs that explain most of

the variation in site-counts across the PGC-1α peaks. The first two components (out of 190 in total) clearly proved to be the most relevant ones, accounting for 10% and 9.6% of the total variation in our dataset, respectively (Fig. 3D). Figure 3E shows the projection of all motifs on these first two principal components, with the names of the motifs with the largest projection indicated in the figure.

Whereas most motifs have projections close to zero along the first component, there is one group of motifs with strong negative projections (ESRRA, NR1H4, NR5A1,2, NR6A1) and one group of motifs with strong positive projections (BACH2, FOS_FOS(B,L1)_JUN(B,D), FOSL2, NFE2, NFE2L1, NFE2L2). These two sets of sites correspond precisely to the two classes of motifs identified above, confirming that the most significant variation in TFBSs across PGC-1α peaks is caused by the occurrence of either ESRRA-like motifs, or FOS-JUN-like motifs. Most interestingly, these two clusters of motifs reflect structurally distinct classes of TFs; the negatively scoring eigenmotifs are characterized by binding of nuclear receptor-type zinc finger domains, while the eigenmotifs with a positive score correspond to activator protein-1 (AP-1)-like leucine zipper domains.

The second principal component corresponds to the strength of the binding signal for these 10 motifs, as confirmed by the robust negative correlation (r=-0.92) between the TFBSs posterior sum per peak and the peak's projection along the second principal component (Fig. 3F).

A

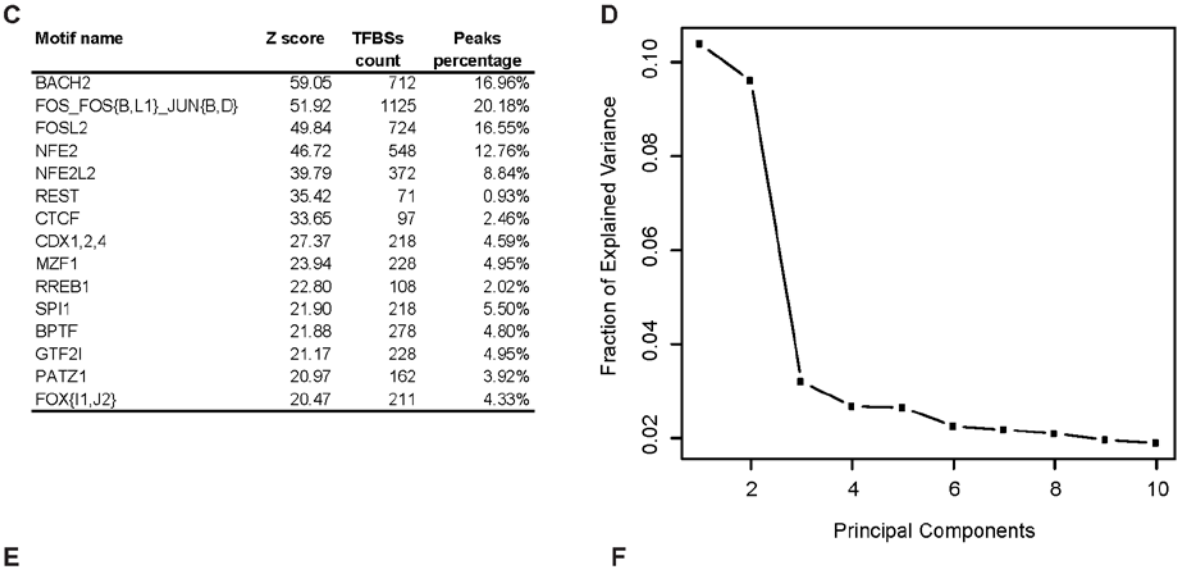


B

Motif name	Z score	TFBSs count	Peaks percentage	Genes in the associated peaks
ESRRA	96.95	1709	19.76%	461 (28.47%)
BACH2	81.74	1300	16.01%	286 (17.67%)
NR5A1,2	80.48	2007	22.58%	522 (32.24%)
NR6A1	72.63	767	8.24%	200 (12.35%)
FOS_FOS{B,L1}_JUN{B,D}	69.89	2030	19.18%	341 (21.06%)
FOSL2	68.06	1360	16.05%	308 (19.02%)
NFE2	62.78	1006	12.15%	229 (14.14%)
NFE2L2	56.66	697	8.53%	174 (10.75%)
HNF4A_NR2F1,2	47.20	766	9.15%	209 (12.91%)
CTCF	47.10	188	2.46%	73 (4.51%)
PPARG	40.61	389	4.30%	101 (6.24%)
REST	40.00	98	0.71%	15 (0.93%)
NR1H4	38.91	586	7.08%	156 (9.64%)
RXRA_VDR{dimer}	35.02	364	4.26%	100 (6.18%)
RREB1	34.23	180	1.78%	42 (2.59%)

C

Motif name	Z score	TFBSs count	Peaks percentage
BACH2	59.05	712	16.96%
FOS_FOS{B,L1}_JUN{B,D}	51.92	1125	20.18%
FOSL2	49.84	724	16.55%
NFE2	46.72	548	12.76%
NFE2L2	39.79	372	8.84%
REST	35.42	71	0.93%
CTCF	33.65	97	2.46%
CDX1,2,4	27.37	218	4.59%
MZF1	23.94	228	4.95%
RREB1	22.80	108	2.02%
SPI1	21.90	218	5.50%
BPTF	21.88	278	4.80%
GTF2I	21.17	228	4.95%
PATZ1	20.97	162	3.92%
FOX{I1,J2}	20.47	211	4.33%



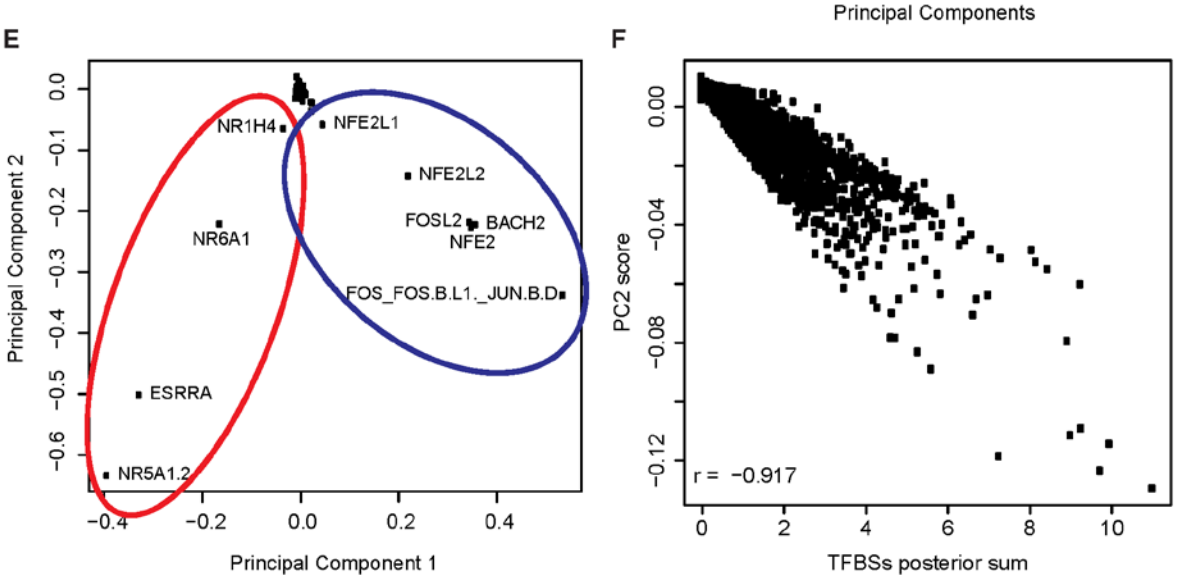


Figure 3. PCA reveals FOS-JUN-like leucine zippers as a new class of putative functional PGC-1α partners. (**A**) Sequence logo of the top position weight matrix discovered *de novo* by PhyloGibbs in the top 200 scoring peaks and of the corresponding canonical motif of ERRα as predicted by STAMP. (**B**) Top scoring results of motif search performed on all 7512 PGC-1α peaks with MotEvo. Motifs depicted in red and blue correspond to the clusters identified by PCA in panel D. (**C**) Top scoring results of motif search

performed on the 3656 “non ESRRA-like” peaks with MotEvo. (**D**) Fraction of explained variance of the top 10 PCA components. (**E**) PCA analysis of the 7512 PGC-1α peaks. Eigenmotif scores across Principal Component 1 (PC1) and Principal Component 2 (PC2) are shown. Red and blue ellipses highlight motif clusters, as identified by PC1, of nuclear hormone receptor-like zinc finger and FOS-JUN-like leucine zipper proteins, respectively. (**F**) Correlation between Principal Component 2 scores and binding site posterior sum for each peak relative to the top 10 PCA motifs. “r” refers to the Pearson correlation coefficient.

4.4.4. Validation of top scoring motifs reveals novel functional partners of PGC-1α

Our analysis identified a number of so-far uncharacterized TFs as potentially functional partners for PGC-1α-controlled gene expression in skeletal muscle cells. In order to experimentally validate some of these candidates, we sorted all TFs by a number of criteria including TFBS over-representation in binding peaks, MARA activity upon PGC-1α over-expression, and the expression pattern of the TFs themselves. Table 1 shows the top 15 ranked TFs according to this selection. As expected, the well-known PGC-1α partner ERRα was identified as the most important factor. For our validation experiments, we chose the next two motifs (FOS_FOS(B,L1)_JUN(B,D) and ZNF143, which is also known as ZFP143) as well as three motifs from further down the list of the top 15 motifs (GTF2I, NFE2L2 and NFYC).

Table 1. Global summary of all analyses performed on PGC-1α peaks. The final score is the count of all analyses where a certain motif passed the defined cutoffs. The motifs chosen for validation and their corresponding values which satisfied the cutoffs are shown in bold.

Motif name	PCA[a]	Over-repr. in all PGC-1α peaks[b]	Over-repr. in “non ESRRA-like” peaks[b]	MARA activity Z score direct[c]	MARA activity Z score indirect[c]	Log2FC in expr. array[d]	Abs. expr. in PGC-1α sample[e]	**Final ranking**
ESRRA	Yes	1	182	6.04(14.78)	15.49(37.94	2.31	1829.45	6
FOS_FOS(B,L1)_JUN(B,D)	**Yes**	**5**	**2**	0.88(2.14)	1.81(-4.34)	**1.78**	**1508.85**	**5**
ZNF143		**27**	**28**	**2.48(6.05)**	**4.65(9.68)**	0.38	**384.36**	**5**
BPTF		21	12	1.38(3.37)	2.56(-6.25)	-0.56	333.34	4
ESR1		17	50	2.33(5.69)	4.53(11.04)	-0.47	232.42	4
FOSL2	Yes	6	3	0.88(2.14)	1.51(3.65)	-0.98	717.09	4
GTF2I		34	**13**	**2.09(5.10)**	**2.38(-5.80)**	-0.55	**1207.81**	**4**
NFE2L2	**Yes**	**8**	**5**	0.57(1.38)	1.01(-2.37)	-0.38	**3673.63**	**4**
NFY(A,B,C)		96	116	**2.37(5.80)**	**3.56(7.62)**	**1.07**	**2409.48**	**4**
NR5A1,2	Yes	3	188	3.53(8.66)	7.73(17.00)	-0.08	80.97	4
REST		12	6	0.48(1.15)	2.41(5.70)	-0.89	328.04	4
RREB1		15	10	1.56(3.82)	2.39(-5.42)	0.05	678.44	4
SP1		24	22	3.99(9.76)	0.61(0.33)	-0.32	751.98	4
STAT2,4,6		29	23	0.35(0.52)	4.81(-9.67)	-2.72	380.12	4
TLX1..3_NFIC(dimer)		19	17	0.84(-2.05)	4.91(-11.97)	-0.34	2339.33	4

[a] Requirement for PCA: being among the top 10 most contributing motifs to PC1 and PC2.
[b] Requirement for motifs over-representation: being among the top 30 significant motifs; ranking position shown.

[c] Requirement for MARA: have a Z-score ≥ 2.0. Numbers between brackets show the difference between the PGC-1α state and the GFP state, representing the direction in which the motif activity changes following PGC-1α over-expression.
[d] Requirement for the expression array (1): having a log2 fold change value ≥ 1.0 (corresponding to 2 folds up-regulation)
[e] Requirement for the expression array (2): having an absolute expression in the PGC-1α sample ≥ 100

FOS, the most up-regulated TF (log2 fold change = 1.78) among the TFs associated with the motif FOS_FOS(B,L1)_JUN(B,D), is a basic leucine zipper transcription factor known to heterodimerize with other leucine zipper proteins in order to form the AP-1 complex [35]. The AP-1 complex furthermore contains JUN as well as ATF proteins. Thus, to dissect the function of the AP-1 protein complex, we also included JUN and ATF3, the most highly expressed isoforms of their respective protein families in muscle cells.

For each of these 7 TFs (ATF3, FOS, GTF2I, JUN, NFE2L2, NFYC and ZFP143), we selected a dozen target genes based on the Chi^2 score of the MARA prediction, presence of a PGC-1α binding peak with at least one predicted binding site for the factor of interest, and at least a 2-fold induction upon over-expression of PGC-1α. As summarized in Fig. 4 and Suppl. Fig. S3, siRNA-based knockdown of all TFs resulted in a robust reduction of the target mRNAs from -40% to -75%. With the exception of NFYC and JUN, we found that the large majority of predicted target genes were down-regulated upon knockdown of the factor, confirming our predictions (Fig. 4). The most consistent effects were observed for FOS and ZFP143 (all targets down-regulated), followed by GTF2I (11 out of 12 down-regulated) and NFE2L2 and ATF3 (10 out of 12 down-regulated). Interestingly, distinct target genes of the AP-1 complex showed differential responsiveness to knockdown of the three AP-1 complex components FOS, JUN and ATF3 (Fig. 4B, Fig. 4C and Fig. 4D). Similarly, PGC-1α-mediated induction of a majority of the predicted target genes for NFE2L2 (Fig. 4E), ZFP143 (Fig. 4F) and GTF2I (Fig. 4G) was reduced upon knockdown of the respective TF when compared to the expression in cells with overexpressed PGC-1α and a scrambled siRNA control. Surprisingly, only 1 of the 11 predicted target genes for NFYC that have been chosen for validation was significantly repressed by siRNA-induced reduction of this TF (Fig. 4H), suggesting that other TFs may be involved in mediating the regulatory effects of the NFYC regulatory motif.

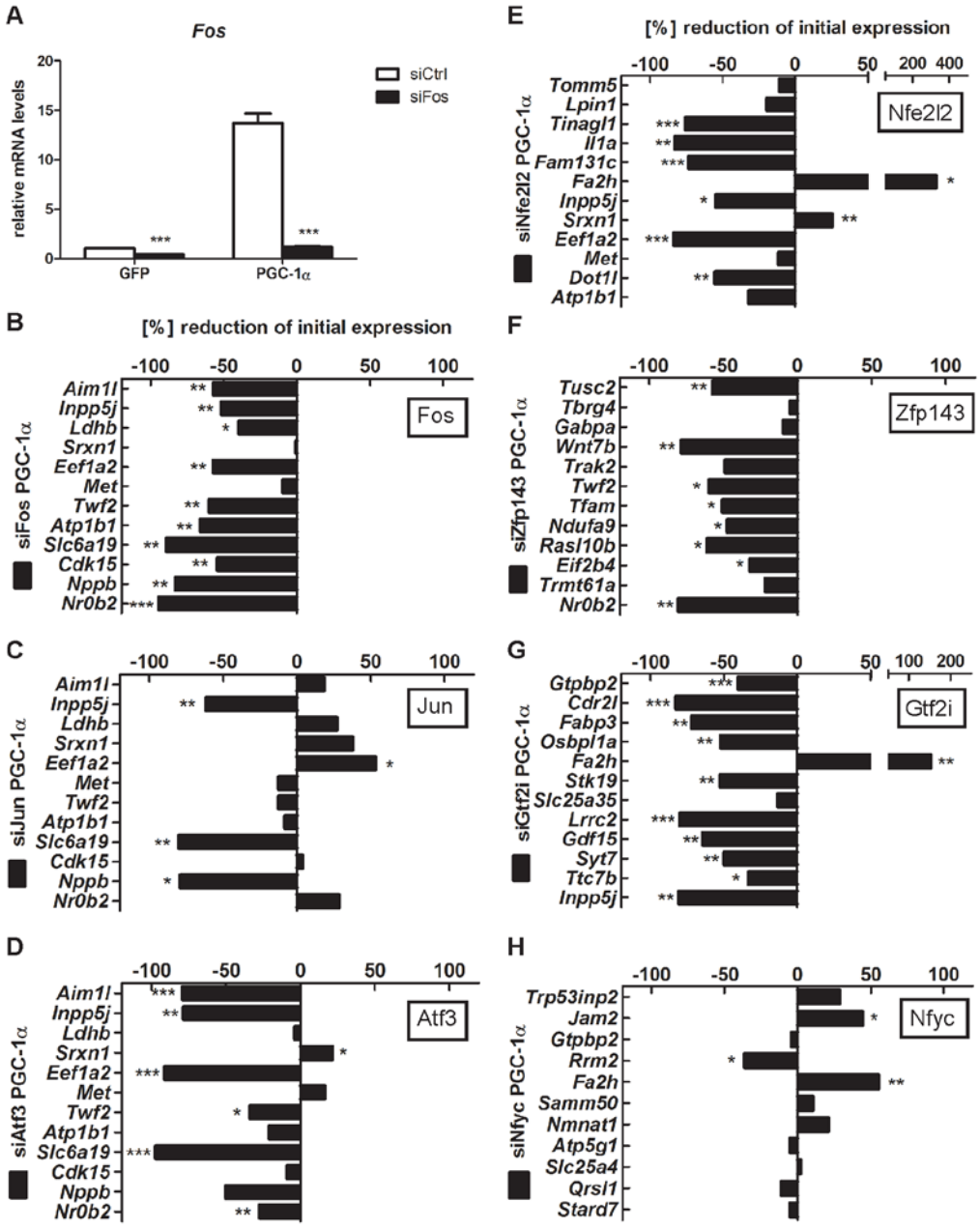


Figure 4. Validation of TFs associated with top scoring motifs reveals novel functional PGC-1α partners. (**A**) siRNA-mediated knockdown efficiency for FOS. Bars represent fold induction over GFP/siCtrl value, error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001. See also Figure S4. (**B-H**) qRT-PCR analysis of PGC-1α target genes whose associated peak contains at least one binding site for the motif: FOS_FOS(B,L1)_JUN(B,D) (**B-D**), NFE2L2 (**E**), ZNF143 (**F**), GTF2I (**G**), NFY(A,B,C) (**H**). Bars represent % change compared to PGC-1α/siCtrl values. Error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

4.4.5. Functional interaction between PGC-1α and different compositions of the AP-1 protein complex

Our targeted validation strategy revealed that PGC-1α target genes predicted to be regulated by the FOS-JUN-like motif react in distinct manners to siRNA-mediated knockdown of individual components of the AP-1 transcription factor protein complex. For example, some genes only reacted to reduction of FOS (Fig. 5A), while others were responsive to the knockdown of two (Fig. 5B) or even all three AP-1 protein partners (Fig. 5C) that we have tested using the siRNA-based approach. To further dissect the responsiveness of PGC-1α target genes to different AP-1 protein complexes, we performed global gene expression arrays upon knockdown of each of the three TF components of the AP-1 complex. Fig. 5D depicts the number of genes that were induced by PGC-1α and that were, at the same time, down-regulated by the siRNA knockdown of any of the three AP-1 complex members. Amongst a total of 477 genes, 89% responded to FOS knockdown, 52% to ATF3 knockdown, and 31% to JUN knockdown. Moreover, while 37% of all targets responded exclusively to FOS, the fraction of targets responding exclusively to either JUN or ATF3 was at most 5%. This analysis shows that, whereas different target genes respond differently to the knockdown of distinct AP-1 components, FOS is the dominant factor in determining AP-1 function in these conditions.

As shown in Fig. 3B, 341 genes were associated to a PGC-1α binding peak containing a predicted site for the FOS-JUN-like motif bound by the AP-1 complex. Of these genes, the expression of 55 was significantly induced by PGC-1α over-expression in muscle cells. In our siRNA-based validation experiment, we found that 47 out of these 55 PGC-1α-induced/AP-1 predicted targets were significantly down-regulated by knockdown of the AP-1 complex components and we called these genes “direct PGC-1α/AP-1 targets”. The remaining 430 genes out of 477 (Fig. 5D) were defined accordingly as “indirect PGC-1α/AP-1 targets” that lack a PGC-1α peak containing a FOS-JUN-like motif, but still are regulated by PGC-1α and the AP-1 protein components (Fig. 5E). To reveal whether these gene categories exert distinct functions, GO and KEGG enrichment analyses were performed. Surprisingly, the 47 direct PGC-1α/AP-1 target genes showed a distinct and significant over-representation of the terms “response to hypoxia” (GO ID: 0001666; adjusted p-value: 0.0247542) and “mTOR signaling pathway” (KEGG ID: mmu04150; adjusted p-value: 0.030674) that were absent in the GO analysis of the remaining PGC-1α/AP-1 targets (Fig. 5F). Recruitment of FOS to the same regulatory regions as PGC-1α in the direct AP-1/PGC-1α target genes was subsequently validated by ChIP (Fig. 5G). These results suggest that AP-1, when interacting with PGC-1α, drives a synergic effect of response to hypoxia; on the other hand, when AP-1 and PGC-1α act separately, and furthermore through downstream intermediate TFs, they regulate the expression of genes involved in mitochondrial organization and energy metabolism.

Intriguingly, several of the predicted AP-1/PGC-1α target genes are also under the control of PGC-1α working with other transcription factors. For example, the vascular endothelial growth factor (VEGF) or, based on the gene expression arrays, 8 OXPHOS genes seem likewise to be under the control of AP-1 as well as ERRα in the context of elevated PGC-1α in skeletal muscle [31, 36]. We therefore assessed the predicted and experimental overlap of these two transcription factors in the regulation of AP-1/PGC-1α target genes. Interestingly, when the PCA analysis of the PGC-1α peaks was stratified in terms of eigenpeaks, we observed two distinct groups of peaks associated with AP-1/PGC-1α target genes (Fig. 5H). First, some of these genes exclusively harbored peaks with FOS-JUN-like TFBSs, whereas the second group exhibited either peaks with both FOS-JUN- and ESRRA-like TFBSs, or a combination of distinct peaks with either of these sites within 10 kb from their promoters (Fig. 5H). Next, we validated this prediction by investigating the change in expression of different AP-1/PGC-1α target genes in the context of reduced ERRα expression and function, elicited by a combination of shRNA-mediated knockdown and pharmacological treatment of muscle cells with the ERRα inverse agonist XCT-790 [31]. In line with the PCA, two distinct groups of ERRα inhibition-sensitive (Fig. 5I-K) and –insensitive (Fig. 5L-N) AP-1/PGC-1α target genes were found.

Finally, since all of the experiments were performed in differentiated myotubes in culture, we assessed whether similar gene expression changes of the direct AP-1/PGC-1α targets involved in hypoxic gene regulation are also observed in skeletal muscle tissue of different gain- [6] and loss-of-function mouse models [7, 8] in vivo. In skeletal muscle-specific PGC-1α knockout mice, the expression of several of these genes was reduced significantly (Fig. 6A-F). Surprisingly however, some of the predicted transcripts were not altered in this loss-of-function model for PGC-1α, for example Nr0b2 (Fig. 6E). To further clarify the role of PGC-1α in the regulation of these genes, relative transcript levels were next assessed in muscle-specific transgenic mice for PGC-1α (Fig. 6G-L). In most cases, the genes with a reduction in their transcription in the PGC-1α muscle-specific knockout animals were inversely elevated in the PGC-1α muscle-specific transgenic mice. Moreover, some of these genes were likewise induced by exercise (Fig. 6G-L) and at least in some cases, for example Twf2 and Nr0b2 (Fig. 6J and K), PGC-1α overexpression and physical activity synergistically boosted gene expression, for Nr0b2even in the absence of any effect of the muscle-specific PGC-1α transgene *per se* (Fig. 6K).

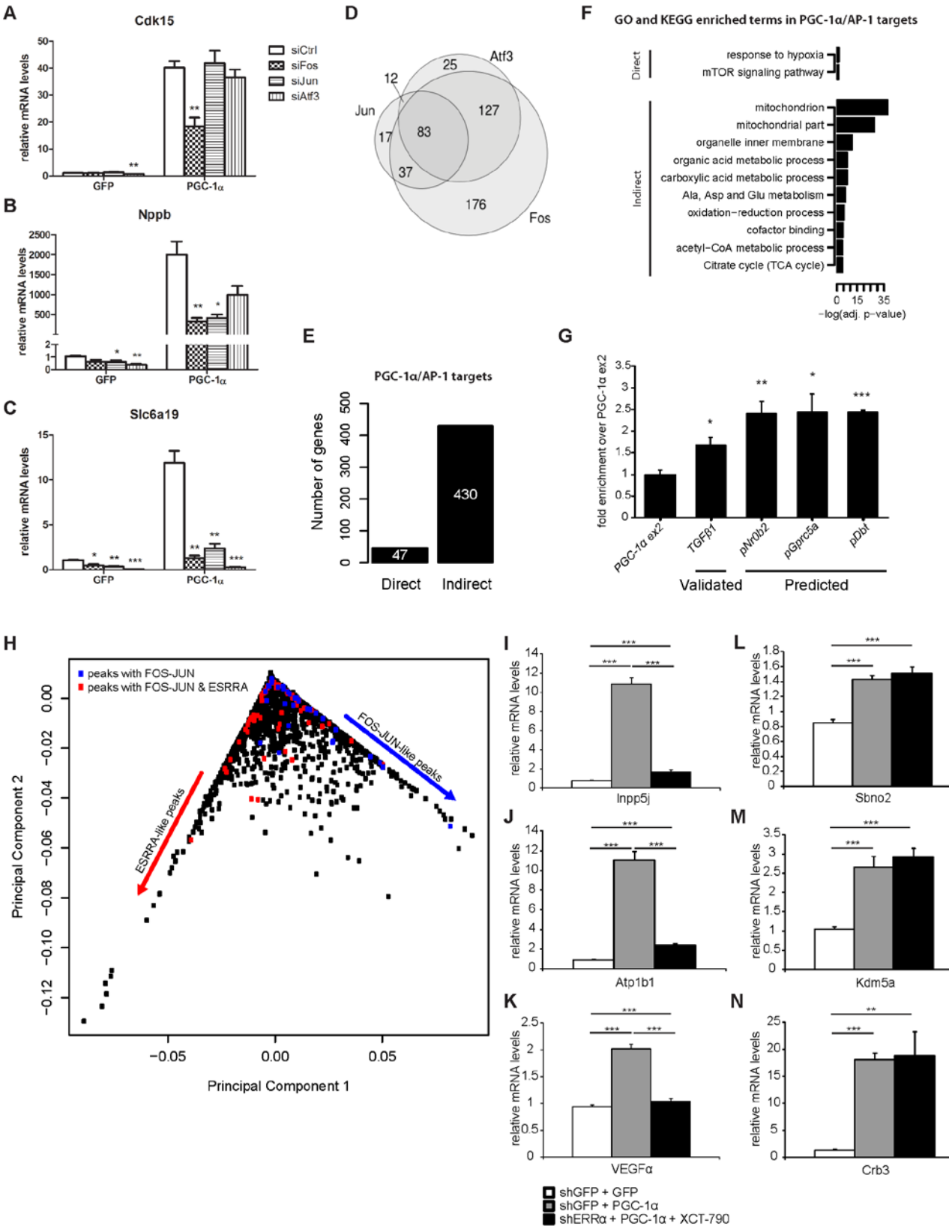


Figure 5. PGC-1α controls the hypoxia gene program via a functional interaction with different configurations of the AP-1 protein complex. (**A-C**) qRT-PCR analysis of *Cdk15*, *Nppb* and *Slc6a19*

mRNA levels in response to PGC-1α over-expression and either siFos, siJun or siAtf3 knockdown. Data are normalized to mRNA levels in GFP infected cells. Error bars represent ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001. (**D**) Venn diagram illustrating the overlap in number of genes up-regulated by PGC-1α and down-regulated by either FOS, JUN or ATF3 knockdown. (**E**) Histogram illustrating the number of direct and indirect PGC-1α/AP-1 target genes. (**F**) Subset of the top significantly enriched Gene Ontology and KEGG terms identified for the two gene groups illustrated in panel (E). (**G**) qRT-PCR validation of the ChIP enrichment of c-Fos measured at the gene of *TGFβ1* (validated) and at the promoters of *Nr0b2*, *Gprc5a* and *Dbt* (predicted) target genes. Bars represent fold enrichment over PGC-1α exon2 set as 1. Error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001. (**H**) PCA analysis of the 7512 PGC-1α peaks. Eigenpeak scores across Principal Component 1 (PC1) and Principal Component 2 (PC2) are shown. Colored dots correspond to peaks associated to the 47 direct PGC-1α/AP-1 targets. Blue dots refer to genes associated to peaks containing only FOS-JUN TFBSs, while red dots refer to genes associated to peaks with FOS-JUN and ESRRA TFBSs, either located in the same peak or in distinct PGC-1α peaks. (**I-K**) qRT-PCR analysis of PGC-1α/AP-1 targets whose associated peaks contain an ESRRA binding site. The bars represent relative mRNA levels compared to AV-shGFP + AV-GFP + vehicle, which is set as 1. The error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001. (**L-N**) qRT-PCR analysis of PGC-1α/AP-1 targets whose associated peaks (if any) do not contain an ESRRA binding site. The bars represent relative mRNA levels compared to AV-shGFP + AV-GFP + vehicle, which is set as 1. The error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001.

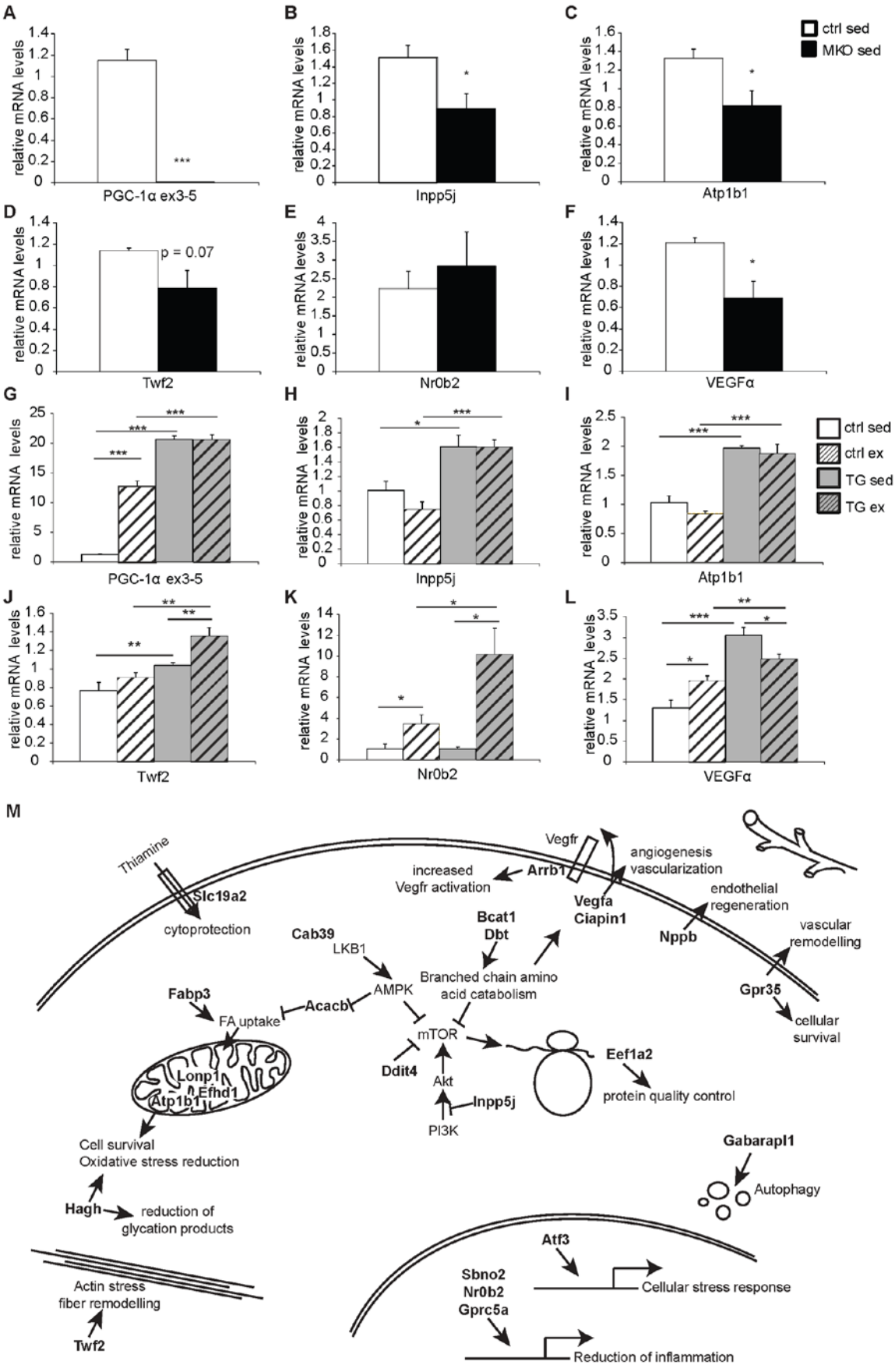

A
B
C
ctrl sed
MKO sed
relative mRNA levels
PGC-1α ex3-5
Inpp5j
Atp1b1
D
E
F
p = 0.07
Twf2
Nr0b2
VEGFα
G
H
I
ctrl ex
TG sed
TG ex
J
K
L
M
Thiamine
Slc19a2
cytoprotection
Vegfr
Arrb1
increased
Vegfr activation
angiogenesis
vascularization
endothelial
regeneration
Vegfa
Ciapin1
Nppb
vascular
remodelling
Cab39
LKB1
AMPK
Bcat1
Dbt
Branched chain amino
acid catabolism
Fabp3
FA uptake
Acacb
Gpr35
cellular
survival
mTOR
Eef1a2
protein quality control
Ddit4
Akt
Inpp5j
PI3K
Lonp1
Efhd1
Atp1b1
Cell survival
Oxidative stress reduction
Hagh
reduction of
glycation products
GabarapI1
Autophagy
Actin stress
fiber remodelling
Twf2
Atf3
Sbno2
Nr0b2
Gprc5a
Cellular stress response
Reduction of inflammation

Figure 6. PGC-1α controls the hypoxic gene program in muscle in vivo. (**A-F**) qRT-PCR analysis of hypoxic genes in sedentary control (ctrl) and muscle-specific knockout mice (MKO). The control group is set as 1. Error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001. (**G-L**) qRT-PCR analysis of hypoxic genes in treadmill running mice. Control (ctrl) and muscle-specific transgenic (TG) mice were used under sedentary and exercise conditions. The control group sedentary is set as 1. Error bars represent SEM. *p < 0.05, **p < 0.01, ***p < 0.001. (**M**) Schematic representation depicting the downstream effects of the functional interaction between PGC-1α and the AP-1 complex in the context of the hypoxia gene program. Direct targets of PGC-1α and AP-1 are indicated in bold.

4.5. Discussion

Exercise-induced skeletal muscle cell plasticity is a highly complex biological program that involves the remodeling of a number of fundamental cellular properties. Since PGC-1α function has been strongly linked to the induction of an endurance-trained muscle phenotype, we here dissected the PGC-1α-controlled transcriptional network in muscle cells. First, our results reveal a broad recruitment of PGC-1α to many different sites in the mouse genome (7512 peaks), the majority of which were either not located within 10 kb distance from a promoter or close to a gene that was not regulated by PGC-1α over-expression at the time of harvest of the cells, as has analogously been observed in many other ChIP-Seq experiments (for example, see ref. [37]). Apart from the fact that PGC-1α could mediate long-range enhancer effects that were excluded in our peak-gene assignment, it is conceivable that PGC-1α recruitment is transcriptionally silent in some binding peaks because it requires the recruitment of additional cofactors for activation, which are not present in the conditions or cell type in which our experiments were performed. In addition, it is possible that a large fraction of PGC-1α binding peaks may be “neutral” in the sense of not having any direct role in regulating gene expression.

Second, while an almost equally strong effect of PGC-1α on gene induction and repression has been reported [31], our analysis now indicates that direct PGC-1α-mediated gene expression is restricted almost exclusively to positively regulated PGC-1α target genes, whereas the vast majority of gene repression is indirect, i.e. not associated with PGC-1α recruitment within a 10 kb distance to their promoters. Thus, the fact that almost 95% of all repressed genes were not linked to PGC-1α recruitment strongly implies that this coregulator primarily acts as a coactivator, and not as a corepressor as suggested by the data of some studies [38-40]. Importantly, indirect repression of PGC-1α target genes was also supported by the MARA prediction. The strong indirect inhibition of genes, many of which are involved in inflammatory processes, is predicted by MARA to be mediated by TFs such as NFκB and IRF factors. Such an indirect inhibition of NFκB and pro-inflammatory genes by PGC-1α in muscle cells has been reported previously [33].

One of the main functions of PGC-1α in all cells and organs is to boost mitochondrial gene transcription and oxidative metabolism. Accordingly, we observed that Gene Ontology terms related to these pathways were highly enriched when analyzing positively regulated PGC-1α target genes in muscle cells. Based on previous studies, the regulation of this core function could have been assigned to the direct interaction of PGC-1α and ERRα binding to regulatory elements of these genes [31, 32]. Surprisingly, our data indicate that many of the genes that are involved in oxidative metabolic pathways are indirectly controlled by PGC-1α and, hence, do not require PGC-1α recruitment to enhancer and promoter elements. Likewise unexpectedly, the MARA analysis implies ERRα action on direct and indirect PGC-1α-induced target genes, i.e. in the presence or absence of PGC-1α coactivation. Thus, while these observations might obviously reflect a temporally distinct control of different PGC-1α target genes that is not represented in our simultaneous analysis of DNA binding and gene expression at one time point, it is conceivable that PGC-1α acts primarily as an upstream regulator of other factors that are subsequently controlling more downstream PGC-1α target genes without direct involvement of PGC-1α itself.

In skeletal muscle, PGC-1α has been reported to interact with ERRs, PPARs and other nuclear receptors, as well as myocyte enhancer and nuclear respiratory factors to mediate transcriptional regulation [3]. Accordingly, ERRα and other nuclear receptor binding motifs were amongst the most highly significant binding elements in our present report. Importantly however, we also predict a number of so-far unknown TFs to functionally interact with PGC-1α and thereby contribute to PGC-1α-controlled gene expression in skeletal muscle. Since a complete functional validation of all new putative TF partners is beyond the scope of this manuscript, we combined the high-throughput results with several computational analyses (see Table 1) to select and test some of the potentially most important factors together with predicted target genes. Notably, in siRNA-based knockdown experiments, we could show that depletion of FOS and its putative AP-1 multimerization partners JUN and ATF3 as well as NFE2L2, ZFP143 and GTF2I in muscle cells reduced the ability of PGC-1α to positively regulate target genes. Second, we could provide evidence of a co-recruitment of FOS and PGC-1α to the same regulatory sites in the vicinity of AP-1/PGC-1α target genes, confirming a functional interaction between these TFs and PGC-1α. Thus, our results indicate that the coactivation repertoire of PGC-1α in muscle exceeds the prediction of previous studies by far. For example, even in our list of the top 15 motifs, several predicted TFs have not yet been investigated in the context of PGC-1α-controlled gene expression, including BPTF, FOSL2, REST or RREB1. Future studies will aim at a more

detailed dissection of the global functional consequences of PGC-1α coactivation of these TFs in muscle cells.

Curiously, almost all of our analyses, and in particular the principal component analysis, highlighted the relevance of FOS-JUN-like motifs. In fact, the largest amount of variation in TFBS occurrence within PGC-1α binding peaks results from either ESRRA-like or FOS-JUN-like motifs. The FOS-JUN-like motif, in particular, embodies the main binding elements of the AP-1 complex, which consists of different configurations of FOS, JUN, ATF and MAF proteins [35, 41]. Our data comparing gene expression in cells with reduced FOS, JUN and ATF3 levels indicate that PGC-1α functionally interacts with the AP-1 complex in different configurations in the regulation of specific genes. The differential requirement observed for distinct AP-1 components might provide an additional layer of control for specific PGC-1α target gene regulation.

AP-1 function itself is regulated by a variety of stimuli, including cytokines, growth factors and stress, and subsequently controls a number of cellular processes including apoptosis, cell proliferation and differentiation, stress response and hypoxia [41, 42]. Mechanistically, we classified PGC-1α-induced/AP-1-knocked-down targets in either direct or indirect genes. Most interestingly, functional analysis of these two groups of genes revealed that when AP-1 and PGC-1α act disjointedly, they are involved in the regulation of mitochondrial and other metabolic genes while, when coactivated by PGC-1α, AP-1 distinctly alters the expression of genes that are enriched in the ontology terms “response to hypoxia” and “mTOR signaling” (Fig. 5F). Intriguingly, a closer analysis of all 47 direct AP-1/PGC-1α target genes revealed 24 genes that are induced by hypoxia, are effectors of hypoxia or attenuate the detrimental consequences of hypoxia (Fig. 6M). For example, several inhibitors of the mTOR signaling pathways are included in this group of genes and hypoxia has been described as a suppressor of mTORC1 activity [43]. Another group of genes contributes to the reduction of cellular stress, detrimental metabolites, reactive oxygen species and increase in cellular survival to reduce potential harmful consequences of prolonged hypoxia [44]. Furthermore, several genes promote endothelial regeneration, vascular remodeling and vascularization [45]. In this context, PGC-1α has previously been shown to promote VEGF-induced angiogenesis in skeletal muscle in a hypoxia-inducible factor 1α (HIF-1α)-independent, ERRα-dependent manner [36]. Similarly, PGC-1α regulates the hypoxic response of brown fat [46], neuronal and endothelial cells [47] even though the mechanisms of cellular protection exerted by PGC-1α in these experimental contexts have not been elucidated. Our findings now indicate that, to ensure adequate oxygen and nutrient supplies for oxidative metabolism in skeletal muscle cells, PGC-1α might coordinate metabolic needs through ERRα-induced *Vegf* expression with a broad, stress-induced AP-1-

dependent hypoxia program. Such a functional convergence was found for a subset of the direct AP-1/PGC-1α target genes that likewise seem to be under the control of ERRα together with PGC-1α (Fig. 5H and I-K). Inversely, for the complementary subset of these genes, the functional interaction between AP-1 and PGC-1α seems distinct from the ERRα-dependent PGC-1α target gene regulation. Finally, in vivo evidence supports our muscle cell cultured-based prediction, considering that many of the AP-1/PGC-1α hypoxia-related target genes exhibit reduced and elevated transcript levels in PGC-1α muscle-specific knockout and transgenic animals, respectively. As previously demonstrated for VEGF and skeletal muscle vascularization [36], many aspects of the phenotypic consequences of exercise-induced muscle hypoxia occur in the muscle-specific transgenic mice even in the absence of physical activity. In extension of these studies, we now however found additional genes involved in this process that show an additional, or in case of Nr0b2, even an exclusive synergistic activation by exercise in the PGC-1α transgenic animals. Thus, combined with previous descriptions of muscle plasticity in these mice post-exercise in regard to insulin sensitivity [29], our present findings reiterate the importance of *bona fide* exercise even in a genetic model for endurance training such as the PGC-1α muscle-specific transgenic animals.

In summary, our data provide a first insight into the transcriptional network controlled by PGC-1α in muscle cells. While one other study of global DNA recruitment of PGC-1α has been performed in the human hepatoma cell line HepG2 [48], our results highlight the importance of combining ChIP-Seq experiment, transcriptional data together with a comprehensive computational modeling approach and experimental validation of predicted key regulators, in order to be able to discover mechanistic as well as functional outcomes of such a network. Combined with the knowledge of transcriptional regulation, posttranslational modifications, alternative splicing and recruitment of different chromatin remodeling protein complexes, a scenario can thus be conceived in which PGC-1α is able to control and integrate different signaling pathways using a multitude of different transcription factor binding partners [10, 11]. A better understanding of such regulatory networks will eventually allow the targeting of whole biological programs or specific sub-modules in pathological states of disregulation.

4.6. Accession numbers

The Gene Expression Omnibus (GEO) accession number for the ChIP-Seq and gene expression array data reported in this paper is GSE51191.

4.7. Acknowledgements

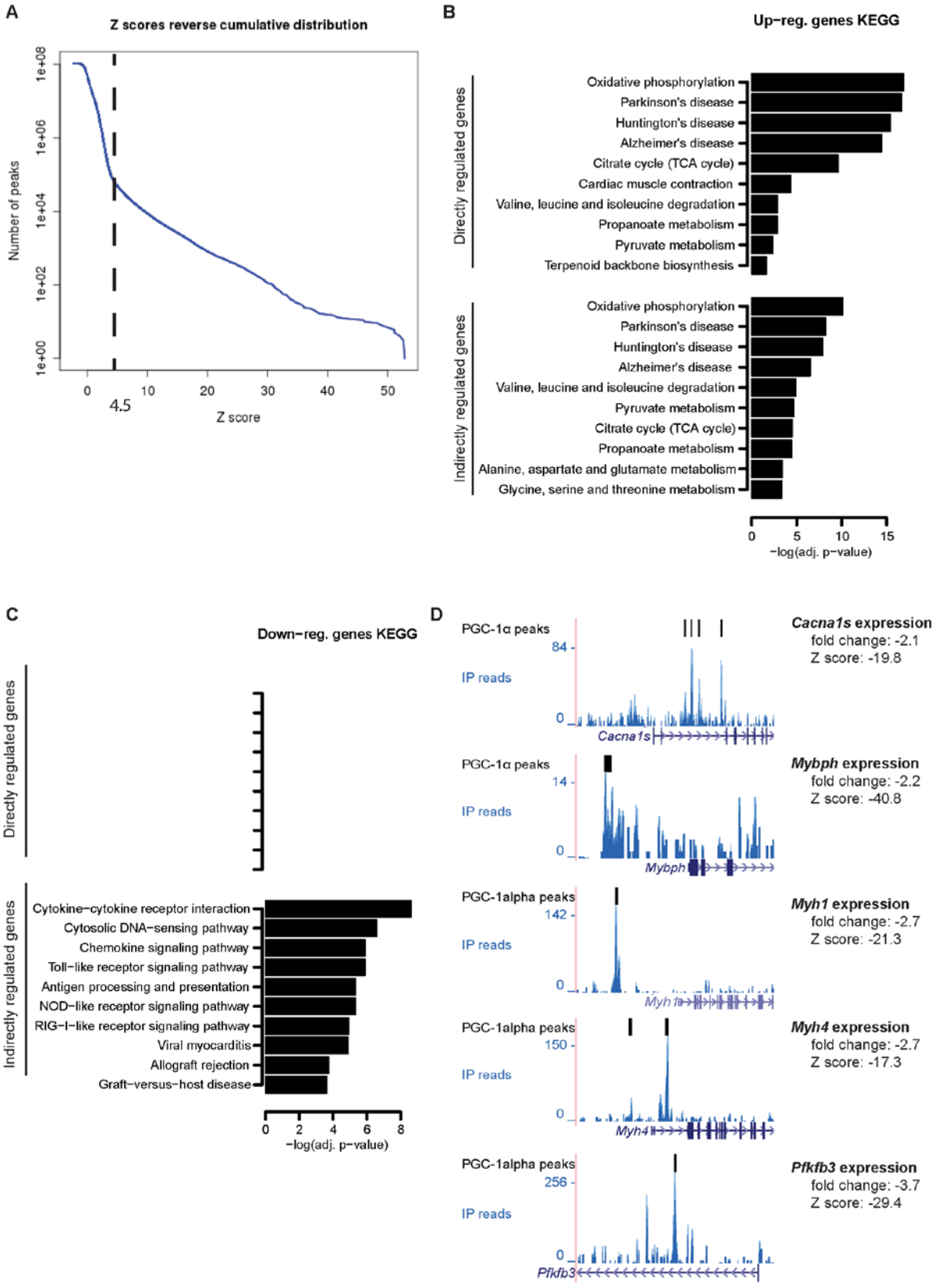
We would like to thank Dr. Anastasia Kralli, Svenia Schnyder, Gesa Santos, Kristoffer Svensson and Markus Beer for reagents, help and input for the preparation of this manuscript. This project was funded by the Swiss National Science Foundation (31003A_135397 to EvN, 310030_132900 to CH), SystemsX.ch (CellPlasticity, StoNets, and BrainstemX research projects to EvN), the Swiss Society for Research on Muscle Diseases (SSEM), the Neuromuscular Research Association Basel (NeRAB), the Gebert-Rüf Foundation "Rare Diseases" Program, the University of Basel and the Biozentrum. SS was supported by an IPhD fellowship of the SystemsX.ch Swiss Initiative in Systems Biology. We are thankful of the [BC]2 Basel Computational Biology Center for providing computational resources.

4.8. Supplementary information

Suppl. Fig. S1. Peak Z score distribution and KEGG functional analysis. Related to Figure 1. (**A**) Distribution of the Z scores for all sliding windows considered by the peak-finding algorithm along the mouse genome. The chosen cutoff for peak calling is depicted by the dotted line. (**B**) Subset of the top significantly enriched KEGG terms identified for direct and indirect up-regulated PGC-1α target genes. (**C**) Subset of the top significantly enriched KEGG terms identified for direct and indirect down-regulated PGC-1α target genes. (**D**) ChIP-Seq signal around the promoter region of the five directly down-regulated genes (*Cacna1s*, *Mybph*, *Myh1*, *Myh4*, *Pfkfb3*) involved in regulating the contractile properties of fast-twitch muscle fibers.

Suppl. Fig. S2. Motif activities clustered by Z score in direct/indirect activation/repression. Related to Figure 2. (**A**) Motifs showing different types of regulation (1=yes, 0=not).

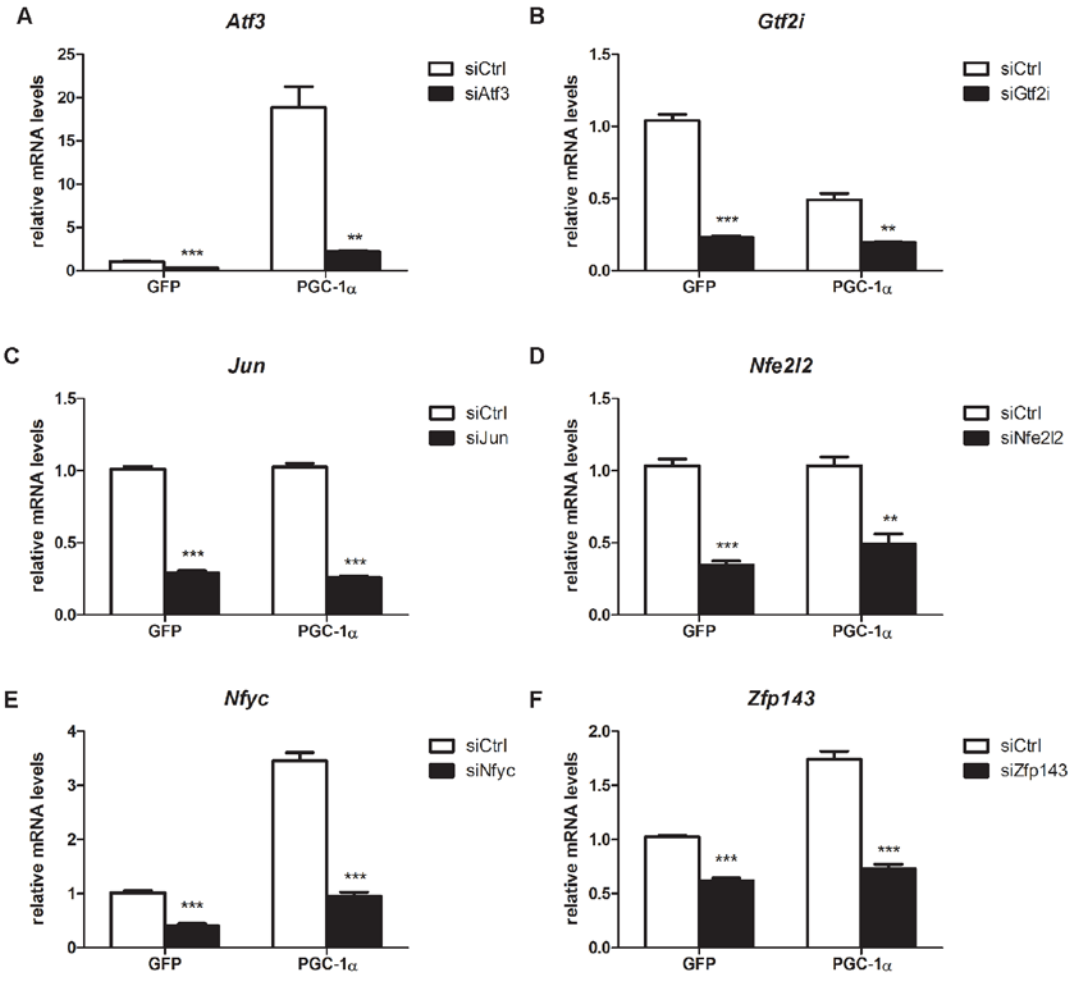
Suppl. Fig. S3. siRNA knockdown efficiency for the putative PGC-1α partner TFs. Related to Figure 4. (**A-F**) siRNA knockdown efficiency for ATF3 (**A**), GTF2I (**B**), JUN (**C**), NFE2L2 (**D**), NFYC (**E**) and ZFP143 (**F**) knockdown. Bars represent fold change over GFP/siCtrl levels. Error bars represent SEM. *p < 0.05; **p < 0.01; ***p < 0.001.


A
Z scores reverse cumulative distribution
Number of peaks
1e+00
1e+02
1e+04
1e+06
1e+08
0
10
20
30
40
50
4.5
Z score
B
Up−reg. genes KEGG
Directly regulated genes
Oxidative phosphorylation
Parkinson's disease
Huntington's disease
Alzheimer's disease
Citrate cycle (TCA cycle)
Cardiac muscle contraction
Valine, leucine and isoleucine degradation
Propanoate metabolism
Pyruvate metabolism
Terpenoid backbone biosynthesis
Indirectly regulated genes
Oxidative phosphorylation
Parkinson's disease
Huntington's disease
Alzheimer's disease
Valine, leucine and isoleucine degradation
Pyruvate metabolism
Citrate cycle (TCA cycle)
Propanoate metabolism
Alanine, aspartate and glutamate metabolism
Glycine, serine and threonine metabolism
0
5
10
15
−log(adj. p−value)
C
Down−reg. genes KEGG
Directly regulated genes
Indirectly regulated genes
Cytokine−cytokine receptor interaction
Cytosolic DNA−sensing pathway
Chemokine signaling pathway
Toll−like receptor signaling pathway
Antigen processing and presentation
NOD−like receptor signaling pathway
RIG−I−like receptor signaling pathway
Viral myocarditis
Allograft rejection
Graft−versus−host disease
0
2
4
6
8
−log(adj. p−value)
D
PGC-1α peaks
IP reads
84
0
Cacna1s
Cacna1s expression
fold change: -2.1
Z score: -19.8
PGC-1α peaks
IP reads
14
0
Mybph
Mybph expression
fold change: -2.2
Z score: -40.8
PGC-1alpha peaks
IP reads
142
0
Myh1
Myh1 expression
fold change: -2.7
Z score: -21.3
PGC-1alpha peaks
IP reads
150
0
Myh4
Myh4 expression
fold change: -2.7
Z score: -17.3
PGC-1alpha peaks
IP reads
256
0
Pfkfb3
Pfkfb3 expression
fold change: -3.7
Z score: -29.4

A

Motif name	Z direct	Z indirect	Z avg. direct	Z avg. indirect	Directly activated	Indirectly activated	Directly repressed	Indirectly repressed
Group 1: motifs only directly activated by PGC-1alpha								
SP1.p2	3.99	0.61	9.76	0.33	1	0	0	0
ELF1,2,4.p2	3.11	1.32	7.59	3.13	1	0	0	0
PAX4.p2	2.50	1.53	6.11	-3.68	1	0	0	0
LMO2.p2	2.36	1.65	5.78	3.98	1	0	0	0
HNF4A_NR2F1,2.p2	2.26	1.54	5.52	3.64	1	0	0	0
GTF2I.p2	2.09	2.38	5.10	-5.80	1	0	0	1
Group 2: motifs directly and indirectly activated by PGC-1alpha								
ESRRA.p2	6.04	15.49	14.78	37.94	1	1	0	0
NR5A1,2.p2	3.53	7.73	8.66	17.00	1	1	0	0
ZNF143.p2	2.48	4.65	6.05	9.68	1	1	0	0
NFY{A,B,C}.p2	2.37	3.56	5.80	7.62	1	1	0	0
ESR1.p2	2.33	4.53	5.69	11.04	1	1	0	0
RXR{A,B,G}.p2	2.29	4.30	5.59	10.50	1	1	0	0
Group 3: motifs only indirectly activated by PGC-1alpha								
NRF1.p2	1.60	4.61	3.91	6.21	0	1	0	0
YY1.p2	0.88	2.97	2.09	5.77	0	1	0	0
EHF.p2	0.73	2.77	1.77	6.35	0	1	0	0
RXRA_VDR{dimer}.p2	0.71	2.54	1.71	6.20	0	1	0	0
HES1.p2	0.34	2.52	0.84	6.10	0	1	0	0
FOXO1,3,4.p2	0.46	2.51	1.13	6.12	0	1	0	0
ELK1,4_GABP{A,B1}.p3	1.18	2.46	2.89	5.95	0	1	0	0
NKX3-1.p2	0.60	2.43	1.48	5.93	0	1	0	0
REST.p3	0.48	2.41	1.15	5.70	0	1	0	0
NFE2L1.p2	1.79	2.32	4.36	5.23	0	1	0	0
POU5F1_SOX2{dimer}.p2	0.24	2.32	0.57	5.65	0	1	0	0
AIRE.p2	0.38	2.24	-0.91	5.40	0	1	0	0
RXRG_dimer.p3	1.67	2.01	4.09	4.89	0	1	0	0
Group 4: motifs only indirectly repressed by PGC-1alpha								
IRF1,2,7.p3	1.77	24.23	4.34	-14.48	0	0	0	1
NFKB1_REL_RELA.p2	0.50	6.54	1.19	-16.01	0	0	0	1
TLX1..3_NFIC{dimer}.p2	0.84	4.91	-2.05	-11.97	0	0	0	1
STAT2,4,6.p2	0.35	4.81	0.52	-9.67	0	0	0	1
DMAP1_NCOR{1,2}_SMARC.p	0.25	4.22	-0.60	-8.73	0	0	0	1
RUNX1..3.p2	0.09	3.94	0.11	-9.61	0	0	0	1
NFATC1..3.p2	0.16	3.46	-0.24	-8.42	0	0	0	1
GATA1..3.p2	1.21	3.39	-2.92	-8.04	0	0	0	1
TBP.p2	1.11	3.20	2.71	-4.04	0	0	0	1
ZIC1..3.p2	0.20	2.99	-0.46	-7.24	0	0	0	1
ATF6.p2	0.24	2.97	-0.51	-7.25	0	0	0	1
TLX2.p2	0.57	2.86	1.37	-6.76	0	0	0	1
TFAP2B.p2	1.75	2.72	4.26	-6.61	0	0	0	1
SPI1.p2	1.69	2.70	4.14	-6.19	0	0	0	1
MEF2{A,B,C,D}.p2	0.97	2.67	2.35	-6.51	0	0	0	1
TFCP2.p2	1.07	2.62	2.57	-5.80	0	0	0	1
BPTF.p2	1.38	2.56	3.37	-6.25	0	0	0	1
LEF1_TCF7_TCF7L1,2.p2	0.17	2.55	0.37	-6.11	0	0	0	1
STAT1,3.p3	0.74	2.53	1.79	-6.17	0	0	0	1
RREB1.p2	1.56	2.39	3.82	-5.42	0	0	0	1
GTF2I.p2	2.09	2.38	5.10	-5.80	1	0	0	1
MYFfamily.p2	0.36	2.38	0.79	-5.12	0	0	0	1
ZNF384.p2	0.64	2.34	-1.55	-5.27	0	0	0	1
TGIF1.p2	0.57	2.34	1.34	-5.68	0	0	0	1
TEAD1.p2	0.99	2.23	-2.43	-5.43	0	0	0	1
SOX{8,9,10}.p2	0.16	2.17	0.30	-5.28	0	0	0	1
CEBPA,B_DDIT3.p2	1.03	2.13	2.51	-5.20	0	0	0	1
MYOD1.p2	1.49	2.05	3.65	-4.99	0	0	0	1

Fig.S3



Suppl. Table 1. Real-time primer sequences. Semiquantitative real-time PCR primers used for validation experiments.

Real-time PCR primers used for testing the efficiency of the ChIP		
Gene promoter or intron	Forward primer	Reverse primer
Tbp intron	TGTGAGCTCCTTGGCTTTTT	ATAGTTGCCCAGCAATCAGG
promoter of *Aco2*	CACCGATAGTTGCTTTCCAGATAC	AACCATCTGACAGGCATAGTCAAT
promoter of *Cycs*	AAGGGCGCCCTCTGGGCACATC	ATCCCCGTCGCGCGCTCACCG
promoter of *Acadm*	CCTTGCCCGAGCCTAAAC	GTCTGGCTGCGCCCTCT
promoter of *Atp5b*	CTGGAAACTTCCACCCTCACTA	GAGAGGTTTTTGGCGGAACTA
promoter of *Idh3a*	GGACGGCGTCAAGGTCAAG	GCCTAGGTGGCCTGTCTGTG

PGC-1α exon 2	TGAGGACCAGCCTCTTTGCCCA	CGCTACACCACTTCAATCCACCC

Gene or gene promoter	Forward + reverse primer		FOS binding site	Peak position
TGFβ1[*1]	F:	TTTGAGACTTTTCCGCTGCT	chr7:26472349-26472356	(see reference [49])
	R:	GGTCCTGCCTCCTTGCGA		
Nr0b2 promoter	F:	GGTACAGCCTGGGTTAATGAC	chr4:133109008-133109015	chr4:133108962-133109162
	R:	ACTGCCTGGATGCCCTTTAT		
Gprc5a promoter	F:	TGATGTCATGAGCCTCACCC	chr6:135011471-135011478	chr6: 135011398-135011598
	R:	TAGCTGTCATTGAGGGCACT		
Dbt promoter	F:	AAGGGGCAAAGCAATTCAGG	chr3: 116215241-116215248 chr3: 116215242-116215249	chr3: 116215152-116215352 chr3: 116215152-116215352
	R:	CTTAGAAAATGTGGTCAGATGCA		

Real-time PCR primers used for testing the knockdown efficiency by siRNAs		
Gene	Forward primer	Reverse primer
Rn18s	AGTCCCTGCCCTTTGTACACA	CGATCCGAGGGCCTCACTA
Fos	TACTACCATTCCCCAGCCGA	GCTGTCACCGTGGGGATAAA
Jun	TGGGCACATCACCACTACAC	TCTGGCTATGCAGTTCAGCC
Atf3	TCTGCGCTGGAGTCAGTTAC	CCGCCTCCTTTTCCTCTCAT
Gtf2i	TTCGAAGGCTTTGCAAGGAAG	TTCGGGGTCCTCACTGGTTT
Nfe2l2	AGTGGATCCGCCAGCTACTC	ATGGGAATGTCTCTGCCAAA
Nfyc	CCACCAGTTCTACGACCACC	GGCCTGTACAATCTGCACCT
Zfp143	GTGGTCGGTCCTTTACCACA	AAATGCCCTCCCACATCCAG

Real-time primers used for target gene validation		
Gene	Forward primer	Reverse primer
Aim1l	CCTGTTGCGTCCATAAGGGT	GCTCTGAGTTCCACATCCCC
Atp1b1	GCTACGAGGCCTACGTGCTA	TGCCACAGTCCTCGAAAATC
Atp5g1	CAGAGGCCCCATCTAAGCAG	TGTCCCGGGAAATGACACTG
Cdk15	ATGCAGTTGCTACCACCGTT	CCGTGGAACTGGATGCTTCT
Cdr2l	GGAACAGGAAAACGAACGGC	ACCACCGTGTACTCACGTTC
Crb3	CCGGACCCTTTCACAAATAGC	CTCTGTCTGCCGCTTTTCC
Dot1l	TGACCTCAGATGAGGAGCCA	TGTCTTCGGGGGAGATTTGC
Eef1a2	CAAGATGGACTCCACGGAAC	CTGGGTTGTAGCCGATCTTC
Eif2b4	ACGGCAAGACCCAATCAGAG	AAGTTCTGCCTTACTCCGGC
Fa2h	GTGGACTGGCAGAAACCTCT	TCTGAGTGGAAGAGGCGAAT
Fabp3	CATGTGCAGAAGTGGAACGG	CTCACCACACTGCCATGAGT
Fam131c	CTGGCTACGTCATCCCTTGT	TCCAGCCTTTCCACTCGAT
Gabpa	GTCGAGGTGGTCATCGATCC	GTAATGTGCTTGGTGCCGTC
Gdf15	CACGCATGCGCAGATCAAAG	TGTGCATAAGAACCACCGGG
Gtpbp2	TGGAAACCTCAAAGCTCGGG	GTACGGAGGGTTGTTGGCTT

Il1a	TGCAAGCTATGGCTCACTTC	GATACTGTCACCCGGCTCTC
Inpp5j	ACAAGGGCGGAGTAAGTGTG	TGAAAGTTATCCTTGCGCTGT
Jam2	GTATTACTGCGAAGCCCGGA	CAACCGTTGCTATGATGCCG
Kdm5a	GTCTTCCGTGTGTCATCAGC	TTAGTCGGGGCAATTCAGGT
Ldhb	GACTCCGAAAATTGTGGCCG	TTCTCTGCACCAGGTTGAGC
Lpin1	CGGCCCTCAACACCAAAAAG	AATTCACCCCACAGCCAGAG
Lrrc2	GTGGAAGGAGCTGCCTGATT	AACAGCTCGATGTACGTGGG
Met	GCTGAGAAACTCTTCCGGCT	AGCCGGCCCATGAATAAGTC
Ndufa9	TTCTGTGGCTCATCCCATCG	TGTAGCCCCAAACACAGTGG
Nmnat1	GGTCGGTGATGCGTACAAGA	CCACGTATCCACTTCCACCC
Nppb	GGCCTCACAAAAGAACACCC	TGCCCAAAGCAGCTTGAGAT
Nr0b2	CCTCTTCAACCCAGATGTGC	GGGCTCCAAGACTTCACACA
Osbpl1a	TCCCCCAATCAGTGCATTCC	GCTTCTACACTCTTGCCCCA
Qrsl1	GTTGGATCAGGGTGCCCTAC	GGGGTTTCTAACTGGCCCAA
Rasl10b	AGACCTGGAAGTGCGGCTAC	GGCAGCGTGCACGTGTTT
Rrm2	TTGCAGCGAGTGATGGCATA	CCATGGCAATTTGGAAGCCA
Samm50	TTTTGATGGACTTGGGCGGA	TGAGATCGCCGCATTACCTC
Sbno2	AGACATCCCAGACACACCTG	TGAGAAGTGGAGTGCTGGAG
Slc25a4	GGTACTTCCCCACTCAAGCC	AGCAAAGTAGCGCCAGAACT
Slc25a35	TAGTCGTGGCAATGACACCC	TCCAAGATCCCCCGGTACAT
Slc6a19	TCCACTCAACCAGAACCAGAC	TGAGTCACTGATGGAAGTGGAG
Srxn1	CCAGGGTGGCGACTACTACT	AGGTCTGAAAGGGTGGACCTC
Stard7	CTCTACGGCCGCCTGTATTC	CGCCATCAAAACAGAGGCAT
Stk19	GTCCTCACTGTCCGAGATGC	CACCATGCTCAGTACAGCCT
Syt7	ACTGGGCAAACGCTACAAGA	TGCAGGCAACTTGATGGCTT
Tbrg4	AACGACAGCCGTACATTGGT	AGCTCCAGGCACTTGTCTTC
Tfam	GAGCGTGCTAAAAGCACTGG	GCTACCCATGCTGGAAAAACA
Tinagl1	TTCTTGTACCAGCGTGGCAT	CCCCACCCAGTGATCTTGAC
Tomm5	CGGAGGAGATGAAGCGGAAG	TATGGAGTGACTCGCAGCAG
Trak2	GCTGAAGAGACGTTCCGCTA	ATCTCGATCCCTCTCTGCCA
Trmt61a	GCTCCTTCTCTCCGTGCATT	TGCGCACATTGTAGACCTGT
Trp53inp2	TACCCCTCCCGCCTGTTTTA	CTGCCGGTGACATAAACGGA
Ttc7b	TGCTCCCCACGATCAAGAAC	ATCTCCCGACTCCTCTCGTC
Tusc2	GCAGTGCCTCCCTTCGTATT	CTGCCCATTCTTGGTGACGA
Twf2	TGCTACCTCCTCTTCCGACT	ATAGCATCTTCAGCCGCACC
VEGFα	CACGACAGAAGGAGAGCAGA	GGGCTTCATCGTTACAGCAG
Wnt7b	TTTCTCTGCTTTGGCGTCCT	GGCCAGGAATCTTGTTGCAG

References

1. Handschin, C. and B.M. Spiegelman, *The role of exercise and PGC1alpha in inflammation and chronic disease.* Nature, 2008. 454(7203): p. 463-9.
2. Pedersen, B.K. and M.A. Febbraio, *Muscles, exercise and obesity: skeletal muscle as a secretory organ.* Nature reviews. Endocrinology, 2012. 8(8): p. 457-65.
3. Handschin, C., *Regulation of skeletal muscle cell plasticity by the peroxisome proliferator-activated receptor gamma coactivator 1alpha.* Journal of receptor and signal transduction research, 2010. 30(6): p. 376-84.
4. Handschin, C. and B.M. Spiegelman, *Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism.* Endocrine reviews, 2006. 27(7): p. 728-35.
5. Finck, B.N. and D.P. Kelly, *PGC-1 coactivators: inducible regulators of energy metabolism in health and disease.* The Journal of clinical investigation, 2006. 116(3): p. 615-22.
6. Lin, J., et al., *Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres.* Nature, 2002. 418(6899): p. 797-801.
7. Handschin, C., et al., *Skeletal muscle fiber-type switching, exercise intolerance, and myopathy in PGC-1alpha muscle-specific knock-out animals.* The Journal of biological chemistry, 2007. 282(41): p. 30014-21.
8. Handschin, C., et al., *Abnormal glucose homeostasis in skeletal muscle-specific PGC-1alpha knockout mice reveals skeletal muscle-pancreatic beta cell crosstalk.* The Journal of clinical investigation, 2007. 117(11): p. 3463-74.
9. Lonard, D.M. and B.W. O'Malley, *The expanding cosmos of nuclear receptor coactivators.* Cell, 2006. 125(3): p. 411-4.
10. Lonard, D.M. and W. O'Malley B, *Nuclear receptor coregulators: judges, juries, and executioners of cellular regulation.* Molecular cell, 2007. 27(5): p. 691-700.
11. Spiegelman, B.M. and R. Heinrich, *Biological control through regulated transcriptional coactivators.* Cell, 2004. 119(2): p. 157-67.
12. Puigserver, P., et al., *Activation of PPARgamma coactivator-1 through transcription factor docking.* Science, 1999. 286(5443): p. 1368-71.
13. Wallberg, A.E., et al., *Coordination of p300-mediated chromatin remodeling and TRAP/mediator function through coactivator PGC-1alpha.* Molecular cell, 2003. 12(5): p. 1137-49.
14. Li, S., et al., *Genome-wide coactivation analysis of PGC-1alpha identifies BAF60a as a regulator of hepatic lipid metabolism.* Cell metabolism, 2008. 8(2): p. 105-17.
15. Lin, J., C. Handschin, and B.M. Spiegelman, *Metabolic control through the PGC-1 family of transcription coactivators.* Cell metabolism, 2005. 1(6): p. 361-70.
16. Handschin, C., *The biology of PGC-1alpha and its therapeutic potential.* Trends in pharmacological sciences, 2009. 30(6): p. 322-9.
17. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome biology, 2009. 10(3): p. R25.
18. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* Journal of molecular biology, 2000. 302(1): p. 205-17.
19. Pachkov, M., et al., *SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates.* Nucleic acids research, 2013. 41(Database issue): p. D214-20.
20. Arnold, P., et al., *MotEvo: integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences.* Bioinformatics, 2012. 28(4): p. 487-94.

21. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics.* Genome biology, 2004. 5(10): p. R80.
22. R Development Core Team, *R: A language and environment for statistical computing.*, 2012, R Foundation for Statistical Computing.
23. Al-Shahrour, F., R. Diaz-Uriarte, and J. Dopazo, *FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.* Bioinformatics, 2004. 20(4): p. 578-80.
24. Suzuki, H., et al., *The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line.* Nature genetics, 2009. 41(5): p. 553-62.
25. Siepel, A., et al., *Evolutionarily conserved elements in vertebrate, insect, worm, and yeast genomes.* Genome research, 2005. 15(8): p. 1034-50.
26. Siddharthan, R., E.D. Siggia, and E. van Nimwegen, *PhyloGibbs: a Gibbs sampling motif finder that incorporates phylogeny.* PLoS computational biology, 2005. 1(7): p. e67.
27. Mahony, S. and P.V. Benos, *STAMP: a web tool for exploring DNA-binding motif similarities.* Nucleic acids research, 2007. 35(Web Server issue): p. W253-8.
28. Summermatter, S., et al., *Skeletal muscle PGC-1alpha controls whole-body lactate homeostasis through estrogen-related receptor alpha-dependent activation of LDH B and repression of LDH A.* Proceedings of the National Academy of Sciences of the United States of America, 2013. 110(21): p. 8738-43.
29. Summermatter, S., et al., *PGC-1alpha improves glucose homeostasis in skeletal muscle in an activity-dependent manner.* Diabetes, 2013. 62(1): p. 85-95.
30. Huss, J.M., et al., *Estrogen-related receptor alpha directs peroxisome proliferator-activated receptor alpha signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle.* Molecular and cellular biology, 2004. 24(20): p. 9079-91.
31. Mootha, V.K., et al., *Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6570-5.
32. Schreiber, S.N., et al., *The estrogen-related receptor alpha (ERRalpha) functions in PPARgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6472-7.
33. Eisele, P.S., et al., *The peroxisome proliferator-activated receptor gamma coactivator 1alpha/beta (PGC-1) coactivators repress the transcriptional activity of NF-kappaB in skeletal muscle cells.* The Journal of biological chemistry, 2013. 288(4): p. 2246-60.
34. Mangelsdorf, D.J. and R.M. Evans, *The RXR heterodimers and orphan receptors.* Cell, 1995. 83(6): p. 841-50.
35. Hai, T. and T. Curran, *Cross-family dimerization of transcription factors Fos/Jun and ATF/CREB alters DNA binding specificity.* Proceedings of the National Academy of Sciences of the United States of America, 1991. 88(9): p. 3720-4.
36. Arany, Z., et al., *HIF-independent regulation of VEGF and angiogenesis by the transcriptional coactivator PGC-1alpha.* Nature, 2008. 451(7181): p. 1008-12.
37. Ma, Z., et al., *Sequence-specific regulator Prdm14 safeguards mouse ESCs from entering extraembryonic endoderm fates.* Nature structural & molecular biology, 2011. 18(2): p. 120-7.
38. Qian, J., et al., *PGC-1alpha regulates hepatic hepcidin expression and iron homeostasis in response to inflammation.* Molecular endocrinology, 2013. 27(4): p. 683-92.
39. Jang, W.G., et al., *Glucocorticoid receptor mediated repression of human insulin gene expression is regulated by PGC-1alpha.* Biochemical and biophysical research communications, 2007. 352(3): p. 716-21.

40. Sandri, M., et al., *PGC-1alpha protects skeletal muscle from atrophy by suppressing FoxO3 action and atrophy-specific gene transcription.* Proceedings of the National Academy of Sciences of the United States of America, 2006. 103(44): p. 16260-5.
41. Shaulian, E. and M. Karin, *AP-1 as a regulator of cell life and death.* Nature cell biology, 2002. 4(5): p. E131-6.
42. Curran, T. and B.R. Franza, Jr., *Fos and Jun: the AP-1 connection.* Cell, 1988. 55(3): p. 395-7.
43. Cam, H., et al., *mTORC1 signaling under hypoxic conditions is controlled by ATM-dependent phosphorylation of HIF-1alpha.* Molecular cell, 2010. 40(4): p. 509-20.
44. Majmundar, A.J., W.J. Wong, and M.C. Simon, *Hypoxia-inducible factors and the response to hypoxic stress.* Molecular cell, 2010. 40(2): p. 294-309.
45. Wagner, P.D., *Skeletal muscle angiogenesis. A possible role for hypoxia.* Advances in experimental medicine and biology, 2001. 502: p. 21-38.
46. Pino, E., et al., *Roles for peroxisome proliferator-activated receptor gamma (PPARgamma) and PPARgamma coactivators 1alpha and 1beta in regulating response of white and brown adipocytes to hypoxia.* The Journal of biological chemistry, 2012. 287(22): p. 18351-8.
47. Zhao, J., et al., *Peroxisome proliferator activated receptor (PPAR)-gamma co-activator 1-alpha and hypoxia induced factor-1alpha mediate neuro- and vascular protection by hypoxic preconditioning in vitro.* Brain research, 2012. 1447: p. 1-8.
48. Charos, A.E., et al., *A highly integrated and complex PPARGC1A transcription factor binding network in HepG2 cells.* Genome research, 2012. 22(9): p. 1668-79.
49. Liu, G., et al., *c-Fos is required for TGFbeta1 production and the associated paracrine migratory effects of human colon carcinoma cells.* Molecular carcinogenesis, 2006. 45(8): p. 582-93.

5. Genome-wide analysis of ERRα in skeletal muscle reveals co-occurrence with SP1 at GC-rich regulatory elements in the absence of direct PGC-1α coactivation

(manuscript currently in preparation)

Silvia Salatino[1,2,3,*], Barbara Kupr[1,*], Mario Baresic[1], Erik van Nimwegen[2,3,#], and Christoph Handschin[1,#]

[1]Focal Area Growth & Development and [2]Focal Area Computational & Systems Biology, Biozentrum, University of Basel, Basel 4056, Switzerland

[3]Swiss Institute of Bioinformatics, Basel 4056, Switzerland

#Correspondence to: Christoph Handschin, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: christoph.handschin@unibas.ch +41 (0)61 267 23 78 or Erik van Nimwegen, Biozentrum, University of Basel, Klingelbergstrasse 70, CH-4056 Basel, Switzerland, email: erik.vannimwegen@unibas.ch +41 (0) 61 267 15 76.

*These authors contributed equally to this manuscript

Declaration of contribution to this study:

- I contributed to the design and organization of the project;

- I performed all the computational analyses of the project, including: raw sequencing data quality filtering and mapping to the reference genome, identification of bound regions, peak annotation, expression array analysis, de novo motif finding and TFBSs over-representation, principal component analysis, gene ontology enrichment analysis, analysis of peaks CG and CpG content, evaluation of motif activity at direct and indirect targets of PGC-1α, analysis of monomer and dimer content of peak subsets;

- I contributed to the analysis and to the interpretation of the results;

- I contributed to the writing of the manuscript and to the creation of the figures.

5.1. Abstract

In skeletal muscle, the estrogen-related receptor α (ERRα) has been shown to play a crucial role in the regulation of various biological processes, which range from glucose and fatty acids metabolism to angiogenesis and reactive oxygen species (ROS) suppression. The peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) is a transcriptional coactivator that controls most of these pathways through direct coactivation of ERRα, which is considered as the major effector of PGC-1α-induced gene expression in the skeletal muscle tissue. However, to which extent ERRα is required for PGC-1α modulation of energy metabolism in skeletal muscle has not been elucidated yet. Also, it is unknown which other transcription factors might co-occur in the vicinity of ERRα sites when there is no direct recruitment of PGC-1α. By combining high-throughput experiments with computational analysis and biochemistry, we mapped the genome-wide recruitment of ERRα to DNA in skeletal muscle and identified the subsets of targets directly or indirectly regulated by this nuclear receptor upon PGC-1α over-expression. Interestingly, we found that ERRα is required for the direct and indirect up-regulation of most of PGC-1α targets. In addition, we show SP1 as partner of ERRα in GC-rich regulatory elements, in the absence of direct coactivation by PGC-1α. Finally, as most of the regulatory effects of PGC-1α on gene regulation were indirect, we focused on the PGC-1α indirectly controlled targets and we were able to identify Nfyc and Znf143 as novel mediators of the indirectly-controlled PGC-1α effects in skeletal muscle.

5.2. Introduction

During the last decade, the transcriptional coactivator peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) has emerged as a major controller of mitochondrial biogenesis and oxidative phosphorylation (OXPHOS) across different tissues and particularly in skeletal muscle [1]. In addition, PGC-1α plays a crucial role in the regulation of angiogenesis, neuromuscular junctions, glucose uptake and fiber type switch [2]. The expression and activity of PGC-1α can be induced and modulated by several physiological stimuli, including cold exposure, fasting and exercise [3]. In skeletal muscle, these signals converge on PGC-1α and are integrated through the coactivation of several transcription factors, particularly nuclear receptors, which in turn act as key regulators of a number of metabolic processes [2].

The transcriptional network controlled by PGC-1α includes, among its members, also the estrogen-related receptor α (ERRα) [4]. In skeletal muscle, ERRα has been shown to play a fundamental role in the control of muscle glucose and fatty acids metabolism [5]; one mechanism responsible for this regulation is triggered by cAMP levels, which act selectively through protein kinase A (PKA) to increase ERRα activity and promote target genes transcription [6]. Moreover, as shown by several studies [4, 7], PGC-1α acts as a potent inducer and coactivator of ERRα, which in turn auto-regulates itself in a positive feedback loop [8]. Although no endogenous ligands have been identified to date for this orphan nuclear receptor, a functional study indicated that ERRα is constitutively active in a ligand-independent manner [9]. Moreover, in differentiated adipocyte cells, the PGC-1α homolog PGC-1β has been suggested to act as a "protein ligand" for the orphan nuclear receptor ERRα [10].

Recently, we have identified ERRα as the major mediator of PGC-1α-induced gene expression in cultured skeletal muscle cells [11]. However, to which extent the induction of PGC-1α target genes requires ERRα is not known. In other words, it is not entirely clear which genes are induced by PGC-1α and ERRα cooperatively, which genes are regulated by ERRα without a direct PGC-1α coactivation and, finally, which genes (if any) are controlled by PGC-1α independently of ERRα. In addition, although a couple of ERRα targets have been identified [8], a complete genome-wide study of ERRα occupancy in skeletal muscle in the context of PGC-1α over-expression has not been conducted so far. By comparing genomic locations occupied by ERRα with those occupied by PGC-1α through chromatin immunoprecipitation followed by deep sequencing (ChIP-Seq), we aim to identify to which regions (and, thus, to which neighboring genes) ERRα and PGC-1α are co-recruited, as opposed to regulatory elements where these two proteins bind independently from each other. Also, we analyze the genomic context which might be responsible for differential binding and investigate whether other transcription factors might be proposed as potential partners of ERRα in the case of an indirect coactivation by PGC-1α.

5.3. Experimental procedures

5.3.1 Cell culture and knockdown of ERRα

C2C12 myoblasts were cultured in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum (FBS), 100 Units/ml penicillin and 100ug/ml streptomycin. After the cells had reached approximately 90% confluence, the differentiation of myoblasts to myotubes was introduced by switching the growth medium to differentiation medium (DMEM supplemented

with 2% horse serum) for 72 hours. For the knockdown and inactivation of ERRα, differentiated C2C12 cells were infected with either the sh-ERRα or sh-GFP (control) adenovirus (AV) and kept in culture for 4 days. These adenoviruses were a generous gift from Prof. A. Kralli from the Scripps Research Institute in La Jolla, California, USA. Subsequently, these cells were infected with either the flag-PGC-1α or GFP adenovirus and kept in culture for two additional days. The differentiation medium, which was used to infect cells, was supplemented with either 2μM of the ERRα inverse agonist XCT-790 (for cells previously infected with the sh-ERRα AV) or with the vehicle (0.02% DMSO, for the remaining cells). The experiments have been performed in biological triplicates. The RNA was isolated using TRIzol® and was performed according to the TRIzol® reagent RNA isolation protocol (Invitrogen).

For the gene expression arrays, the RNAs from the following three conditions were used: shGFP AV + GFP AV + vehicle (0.02% DMSO); shGFP AV + flag-PGC-1α AV + vehicle (0.02% DMSO); shERRα AV + flag-PGC-1α AV + 2μM XCT-790.

5.3.2. ChIP and ChIP-Seq

Chromatin immunoprecipitation (ChIP) was performed according to the Agilent Mammalian ChIP-on-chip Protocol version 10.0. For every immunoprecipitation, we used approximately $1x10^8$ C2C12 cells, which were differentiated to myotubes and infected with flag-PGC-1α adenovirus. For the cross-linking, the cells were incubated in a 1% formaldehyde solution for 10 minutes. After the lysis and sonication but before the immunoprecipitation (IP), 50μl of the lysate was saved for the extraction of input DNA. For the immunoprecipitation of ERRα, the magnetic beads (Dynabeads® Protein G, Invitrogen) were coated with the monoclonal anti-ERRα antibody (ERRα Rabbit Monoclonal Antibody, Clone ID: EPR46Y, Epitomics). For the ChIP of SP1, the magnetic beads (Dynabeads® Protein G, Invitrogen) were coated with the polyclonal anti-SP1 antibody (ChIPAb+™Sp1 Rabbit Polyclonal Antibody, #17-601, Millipore). The ChIP was performed overnight on a rotating platform at 4°C. The PGC-1α ChIP-Seq data was obtained from previous experiments and was performed using the monoclonal anti-flag antibodies (Monoclonal ANTI-FLAG® M2 Antibody, Sigma) for the immunoprecipitation [11].

5.3.3. High-throughput sequencing and read mapping

The ERRα ChIP-Seq experiment in C2C12 cells undergoing PGC-1α over-expression was performed at the joint Quantitative Genomics core facility of the University of Basel and the Department of Biosystems Science and Engineering (D-BSSE) of the ETH Zurich in Basel. The DNA libraries for ERRα and the whole cell extract (WCE) were prepared using the standard

Illumina ChIP-Seq protocol, as described by the manufacturer. The immunoprecipitated samples were sequenced on the Illumina® HiSeq2000 sequencer.

The sequenced reads underwent a quality filter which discarded all reads having Phred score >= 20, read length >= 25 bps and ambiguous nucleotides (Ns) per read <= 2. The reads that passed the filter were used as input for Bowtie version 0.12.7 [12] and aligned to the UCSC mm9 mouse genome assembly. Moreover, to avoid any PCR amplification error which might have arisen during sample preparation, we removed redundant reads mapping to the same location with the same orientation and we kept only one read per position. Consequently, we obtained 2'155'507 reads for the IP and 84'175'472 reads for the WCE.

5.3.4. Peak calling

To detect regions that were significantly enriched in ERRα binding, we estimated the fraction of all ChIP reads f_{IP} that fell within consecutive 200 bps sliding windows genome-wide and, in parallel, the fraction f_{WCE} of reads from the whole cell extract that fell in a 2000 bps long window centered on the same position of the IP window. Using these two sets of values, we quantified the ChIP enrichment of each window as:

$$Z = \frac{f_{IP} - f_{WCE}}{\sqrt{{\sigma^2}_{IP} + {\sigma^2}_{WCE}}}$$

where ${\sigma^2}_{IP}$ and ${\sigma^2}_{WCE}$ are the IP and WCE read frequency variances, given by:

$${\sigma^2}_{IP} = \frac{f_{IP} * (1 - f_{IP})}{N_{IP}} \text{ and } {\sigma^2}_{WCE} = \frac{f_{WCE} * (1 - f_{WCE})}{N_{WCE}}$$

respectively. All consecutive windows having a Z-score greater than 3.5 were merged and the top scoring one from each window cluster was considered as the peak summit and used for further analyses.

5.3.5. Transcription factor binding sites over-representation

In order to account for the conservation of transcription factor binding sites (TFBSs) across related species, we aligned the ERRα binding peaks to their orthologous regions from 6 mammalian species – human (hg18), opossum (monDom4), dog (canFam2), rhesus macaque (rheMac2), horse (equCab1) and cow (bosTau3) – through the software T-Coffee [13]. Using the MotEvo algorithm [14] on these alignments, we predicted binding sites within our peaks for a

collection of 190 mammalian regulatory motifs (also called weight matrices, or WMs) that was downloaded from the SwissRegulon database [15].

To assess the over-representation of the predicted binding sites (i.e. those with a posterior probability greater than 0.1) within the ERRα peaks, we created a set of background regions by shuffling the columns of the multiple alignments and maintaining, at the same time, both the conservation pattern and the gap patterns of the original aligned nucleotides. MotEvo was run on the shuffled alignments using the same settings as for the original sequences. Significant enrichment of binding sites for each motif *x* in the ERRα peaks was computed by collecting the sum n_x of the posterior probabilities for its predicted TFBSs along ERRα peak alignments and the corresponding sum n'_x in the shuffled alignments; we quantified the motif over-representation as:

$$Z = \frac{f_x - f'_x}{\sqrt{\frac{f_x * (1 - f_x)}{L_x} + \frac{f'_x * (1 - f'_x)}{L'_x}}}$$

where L_x and L'_x are the total lengths of the original and shuffled alignments, respectively, while f_x and f'_x are given by the equations:

$$n_x * l_x = f_x * L_x \text{ and } n'_x * l_x = f'_x * L'_x$$

with l_x the length of motif *x*.

5.3.6. Principal Component Analysis of TFBS occurrences within peaks

A Principal Component Analysis (PCA) was performed using the "svd" package of the R programming language [16]. The input data for the PCA was a *p* × *m* matrix *N*, composed by the total number of binding sites N_{pm} predicted for each of the 3225 ERRα binding peaks *p* (rows) using our collection of 190 mammalian regulatory motifs *m* (columns). Singular Value Decomposition (SVD) was applied to the mean centered matrix $\widetilde{N}_{pm} = N_{pm} - \langle N_m \rangle$ using the 'svd' package of the statistical software R.

5.3.7. Gene expression array analysis

The .CEL files containing the raw probe intensities were processed using the Bioconductor package 'affy' [17], which corrects for background and unspecific binding. After using the 'Mclust' R package [16] to distinguish expressed from non-expressed probes, the intensity values were

quantile normalized across all samples and mapped to the mm9 mouse transcripts UCSC collection. These were further associated to a comprehensive collection of mouse promoters that was downloaded from the SwissRegulon database [15]. For each promoter, the log2 fold change (log2FC) was compared between the following conditions: over-expressed PGC-1α (treatment) and GFP (control); ERRα knockdown with the addition of XCT-790 (treatment) and over-expressed PGC-1α (control). The significance of the expression change was assessed by a Z score, which was computed as:

$$Z = \frac{\bar{E}_{treatment} - \bar{E}_{control}}{\sqrt{\frac{{\sigma^2}_{treatment}}{n} + \frac{{\sigma^2}_{control}}{n}}}$$

where $n = 3$ was the number of replicate samples, $\bar{E}_{treatment}$ is the mean log2 expression across the treatment samples, $\bar{E}_{control}$ is the mean log2 expression across the control samples, and ${\sigma^2}_{treatment}$ and ${\sigma^2}_{control}$ are the variances of log2 expression levels across the replicates for the treatment and control samples, respectively. A log2FC threshold of ±0.5849625 (corresponding, in a more commonly used notation, to 1.5 fold change, when taking the inverse of the linear binding ratio) and a Z score cutoff of ±3 were used to identify significantly up-/down-regulated promoters. The criterion adopted to assign peaks to promoters was proximity. Each peak was assigned to its closest promoter (and, thus, to its associated gene symbol) whenever the distance between their central positions was smaller than 10 kb; otherwise, the peak was not assigned to any promoter. Hence, a peak could have been assigned to a promoter also in the case it was located within an intron or an exon.

5.3.8. Ontology terms over-representation analysis

Gene Ontology analysis was performed using the functional analysis software FatiGO [18]. Only ontology terms having a false discovery rate (FDR) adjusted p-value <= 0.05 for enrichment were considered significant.

5.3.9. Motif activity response analysis

The information obtained from the sequencing data and from the microarrays were integrated by extending ISMARA [19] in order to model the direct and indirect regulatory effects that ERRα and PGC-1α exert on their downstream targets. Using the input expression data and the computationally predicted transcription factor binding sites, ISMARA infers, for each of 190 regulatory motifs *m* (associated with roughly 350 mouse TFs), the activity A_{ms} of the motif in

each sample *s* when the motif occurs *outside* of a binding peak, and the activity A^*_{ms} of the motif when it occurs *within* a binding peak of either PGC-1α, ERRα or both (overlapping peaks). A significant change in the motif activity A_{ms} upon knockdown of ERRα and/or over-expression of PGC-1α indicates an indirect regulatory effect of either ERRα, PGC-1α or both of them on the motif *m*, whereas a change in the motif activity A^*_{ms} reflects a direct regulatory effect as mediated by the motif *m*. For each promoter *p* that was not associated with any binding peak (which we denote as "indirect target"), we modeled its log-expression in sample *s,* e_{ps}, in terms of the predicted number of TFBSs N_{pm} that occur in the proximal promoter region (running from -500 to +500 relative to the transcription start site, orTSS) for each regulatory motif *m.* That is, ISMARA assumes the linear model:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N_{pm} A_{ms}$$

where c_p is the basal expression of promoter *p*, $\tilde{c}_s$ is a sample-dependent normalization constant, and A_{ms} is the regulatory activity of motif *m* in sample *s*, which is inferred by the model. Formally, A_{ms} quantifies the amount by which the expression of promoter *p* in sample *s* would be reduced if a binding site for motif *m* were to be deleted from the promoter. For each "direct target" promoter *p* that has an associated binding peak (which could be an ERRα, a PGC-1α or an overlapping ERRα/PGC-1α peak) within 10 kb, we modeled its expression in terms of the predicted TFBSs in the binding peak, i.e.:

$$e_{ps} = c_p + \tilde{c}_s + \sum_m N^*_{pm} A^*_{ms}$$

where N^*_{pm} is the number of predicted TFBSs for motif *m* in the peak associated with promoter *p*, and A^*_{ms} is the motif activity of regulator *m* in sample *s* when this motif occurs in the context of either ERRα binding, PGC-1α recruitment or both. Besides motif activities MARA also calculates error-bars δ_{ms} for each motif *m* in each sample *s*. Using these, MARA calculates, for each motif *m*, an overall significance measure for the variation in motif activities across the samples analogous to a z-statistic:

$$z_m = \sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\frac{A_{ms}}{\delta_{ms}}\right)^2}$$

For each motif we calculate a z-score Z_m associated with its indirect activity changes, a z-score $Z^*_{m,ERR\alpha}$ associated with its direct activity changes in the context of ERRα binding, a z-score $Z^*_{m,PGC1\alpha}$ associated with its direct activity changes in the context of PGC-1α recruitment and a z-score $Z^*_{m,BOTH}$ associated with its direct activity changes in the context of both ERRα binding and PGC-1α recruitment.

5.3.10. Quantitative real-time PCR and statistical analysis

Quantitative real-time PCR (qRT-PCR) was used to validate the efficiency of the ERRα knockdown and to verify that the ChIP of ERRα and the ChIP of SP1 were successful. The sequences of all primers, which were used for qRT-PCR, are listed in Suppl. Table 1. Regarding the statistical analysis of qRT-PCR data sets, the values are presented as the mean ±SEM. Student's t tests were performed and a p-value < 0.05 was considered as significant. *<0.05, **<0.01, ***<0.001

5.4. Results

5.4.1. ERRα can be recruited to DNA also independently from PGC-1α

The capability of ERRα to control the transcription of its downstream target genes has been postulated to be strictly dependent on the coactivation of PGC-1α, to the extent that this co-regulator – more precisely its homolog PGC-1β – has been proposed to act as a "protein ligand" for ERRα [10]. To verify this assumption and to identify all regions genome-wide that are bound by this transcription factor in skeletal muscle, we performed a chromatin immunoprecipitation (ChIP) experiment followed by high-throughput sequencing (ChIP-Seq) of ERRα in differentiated C2C12 murine myotubes. The identified ERRα binding sites were then compared to a previously sequenced set of PGC-1α recruitment regions [11] obtained within the same context of adenovirus-induced PGC-1α over-expression. In order to identify all genomic locations significantly enriched in ERRα binding, we passed a sliding window along the genome and compared, for each consecutive window, the local IP read density with the whole cell extract (WCE) background read amount. All regions with a Z score bigger than 3.5 were merged in a final total of 3225 peaks, which included binding regions in the vicinity of known ERRα target genes (Suppl. Fig. S1A), like the pyruvate dehydrogenase lipoamide kinase isozyme 4 (Pdk4) and the isocitrate dehydrogenase 3 [NAD+] alpha (Idh3a) [20, 21]. The enrichment of IP

fragments from the ChIP-Seq experiment was validated for some of these ERRα target genes by semiquantitative real-time PCR (Suppl. Fig. S1B).

Most surprisingly, when we compared ERRα and PGC-1α occurrences genome-wide (Fig. 1A), we noticed that the majority of ERRα (~60%) peaks are actually not overlapping a PGC-1α peak, suggesting that the so-far believed concept of symbiotic cooperation between these two proteins is in fact only restricted to a small percentage of their targets (~40% for ERRα and ~18% for PGC-1α), or that all the other ERRα peaks are non-functional. Some examples of this differential regulation are depicted in Fig. 1B. Moreover, of the 1321 ERRα peaks overlapping a PGC-1α site (that is, they share at least one base pair), the vast majority of them are well centered on the closest PGC-1α peak at a distance of a couple of dozen base pairs (Fig. 1C). Noticeably, a big fraction of ERRα peaks resides within 100 bp from the closest mouse promoter region (Fig. 1D), as opposed from the PGC-1α peaks which we previously found to be more distally located.

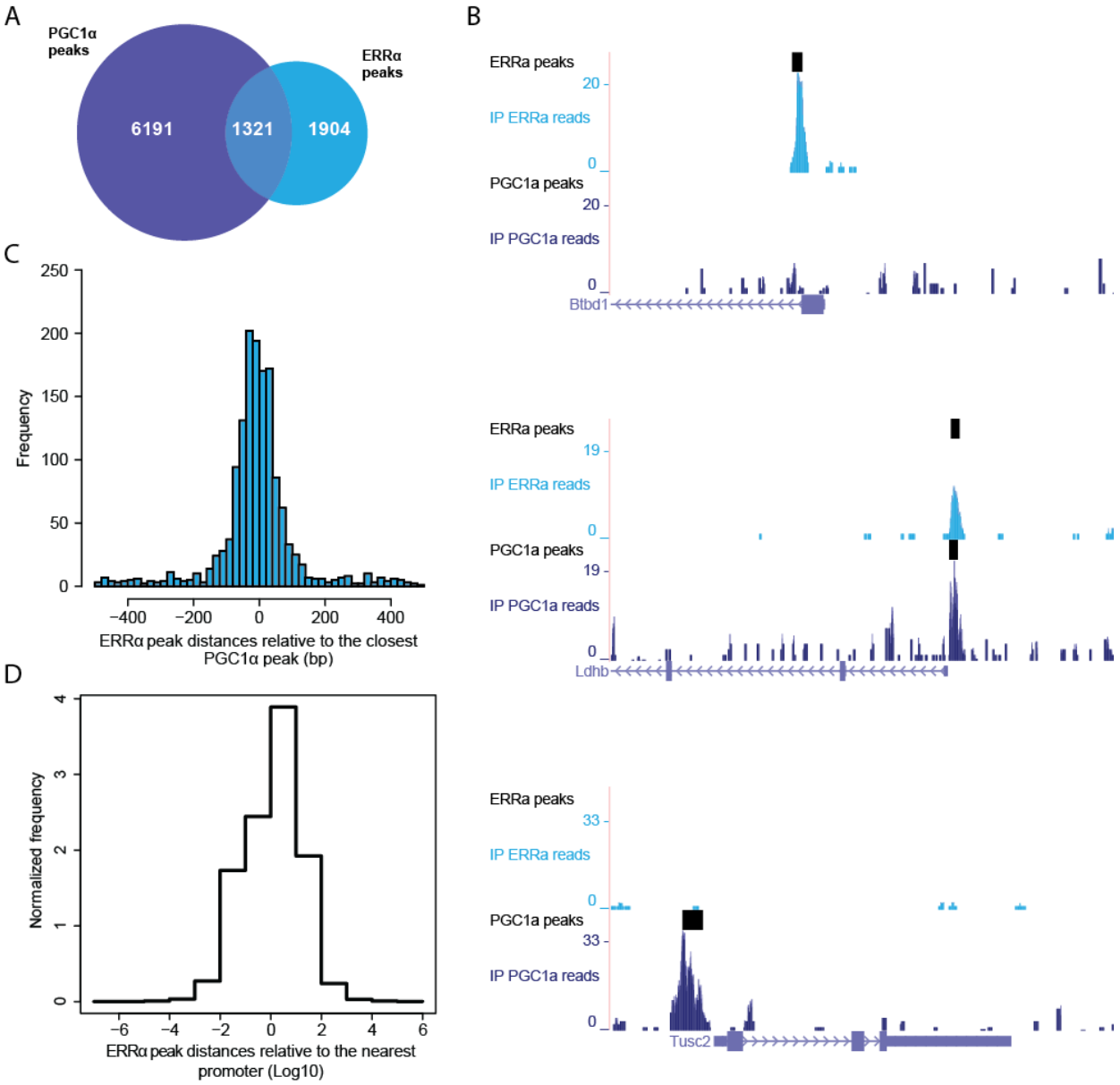


Figure 1. ERRα and PGC-1α are recruited to both shared and distinct sets of target genes. (**A**) Venn diagram depicting the number of ChIP-Seq binding peaks for PGC-1α (blue) and for ERRα (cyan). (**B**) PGC-1α and ERRα read densities around the TSS of the genes Btbd1, Ldhb and Tusc2 obtained from the UCSC Genome Browser. The first plot is an example of a gene directly regulated by ERRα, but not by PGC-1α; the second is directly regulated by both ERRα and PGC-1α; the third is directly regulated by PGC-1α, but not by ERRα. (**C**) Distribution of ERRα peaks relative to their closest PGC-1α peaks. (**D**) Distribution of ERRα peaks from the nearest mouse promoter region.

5.4.2. PGC-1α induces gene transcription both dependently and independently on ERRα

To integrate the results obtained from the ChIP-Seq experiment with functional data, we further analyzed the impact of ERRα on its downstream targets in terms of their gene expression

patterns in differentiated muscle cells under the following conditions: (i) shGFP-transfected cells over-expressing GFP; (ii) shGFP-transfected cells over-expressing PGC-1α; (iii) shERRα-transfected cells over-expressing PGC-1α and supplemented with the ERRα inverse agonist XCT-790. By comparing the second with the first and the third with the second condition, we sought to assess both the effect of PGC-1α over-expression and the response to ERRα knockdown (Fig. 2A). After mapping the microarray probes to known transcripts and, through these, to a reference set of mouse promoters [15], we noticed that more promoters were significantly up-regulated (1863, corresponding to 1165 genes) than down-regulated (658, corresponding to 469 genes) following PGC-1α over-expression; in contrast, following ERRα knockdown, we observed the opposite effect: 910 promoters (corresponding to 597 genes) were significantly induced whereas 1952 promoters (corresponding to 1205 genes) were repressed (Fig. 2B).

By combining the expression array with ERRα and PGC-1α genomic occupancy data, we quantified and distinguished PGC-1α directly or indirectly up-regulated target genes in two categories: ERRα-dependent and ERRα-independent targets. In particular, all array genes were defined ERRα-dependent when they were significantly repressed upon ERRα knockdown (compared to the PGC-1α over-expression condition) and ERRα-independent in the remaining cases. In the case of directly up-regulated PGC-1α targets, we noticed that ~70% of them were ERRα-dependent (Fig. 2C); similarly, ~62% of indirectly up-regulated PGC-1α targets were ERRα-dependent (Fig. 2D). We furthermore used the ERRα-dependent and -independent PGC-1α direct or indirect target genes as input for the functional annotation tool FatiGO [18] to stratify these in terms of the regulatory pathways involved (Figs. 2E and 2F). As expected, most of the enriched terms were related to mitochondria and oxidative energy metabolism. However, we obtained more significantly enriched terms for ERRα-dependent than for ERRα-independent genes, suggesting that most of the biological pathways controlled by PGC-1α in skeletal muscle cells require ERRα. Interestingly, the top gene annotations were very similar among ERRα-dependent and ERRα-independent PGC-1α targets, suggesting that the biological role of PGC-1α in the modulation of mitochondrial function and organization is partly regulated with ERRα, but partly also without. As shown in Fig. 2G and Fig. 2H, we confirmed by qRT-PCR two ERRα-dependent (Aim1l and Twf2) and two ERRα-independent (Atg9b and Ifrd1) PGC-1α target genes.

For completeness, we also checked genes dependency on ERRα for PGC-1α down-regulated targets. Since PGC-1α is a coactivator rather than a corepressor, we assumed that the 22 directly down-regulated genes might represent spurious peak-to-gene associations. In fact, the

vast majority of down-regulated genes (~95%) were indirect PGC-1α targets, confirming the results we previously found [11]. Out of these 447 indirectly down-regulated genes, about a quarter (104 targets) were dependent on ERRα – that is, significantly induced upon ERRα knockdown –, whereas the remaining 343 were not.

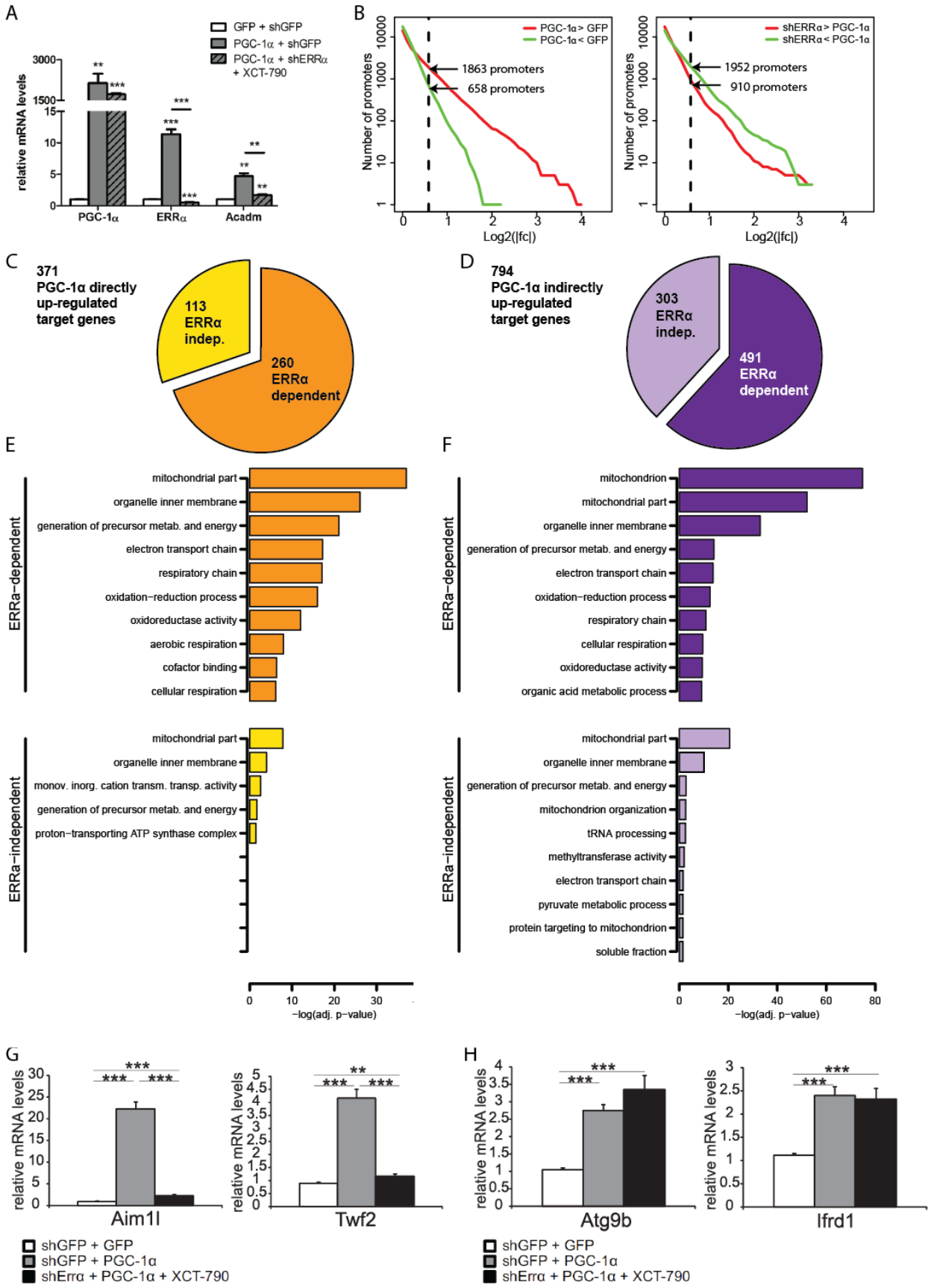

A
GFP + shGFP
PGC-1α + shGFP
PGC-1α + shERRα + XCT-790
relative mRNA levels
PGC-1α
ERRα
Acadm
B
PGC-1α> GFP
PGC-1α< GFP
1863 promoters
658 promoters
Number of promoters
Log2(|fc|)
shERRα> PGC-1α
shERRα< PGC-1α
1952 promoters
910 promoters
C
371
PGC-1α directly up-regulated target genes
113 ERRα indep.
260 ERRα dependent
D
794
PGC-1α indirectly up-regulated target genes
303 ERRα indep.
491 ERRα dependent
E
ERRα−dependent
mitochondrial part
organelle inner membrane
generation of precursor metab. and energy
electron transport chain
respiratory chain
oxidation−reduction process
oxidoreductase activity
aerobic respiration
cofactor binding
cellular respiration
ERRα−independent
mitochondrial part
organelle inner membrane
monov. inorg. cation transm. transp. activity
generation of precursor metab. and energy
proton−transporting ATP synthase complex
−log(adj. p−value)
F
ERRα−dependent
mitochondrion
mitochondrial part
organelle inner membrane
generation of precursor metab. and energy
electron transport chain
oxidation−reduction process
respiratory chain
cellular respiration
oxidoreductase activity
organic acid metabolic process
ERRα−independent
mitochondrial part
organelle inner membrane
generation of precursor metab. and energy
mitochondrion organization
tRNA processing
methyltransferase activity
electron transport chain
pyruvate metabolic process
protein targeting to mitochondrion
soluble fraction
−log(adj. p−value)
G
relative mRNA levels
Aim1l
Twf2
shGFP + GFP
shGFP + PGC-1α
shErrα + PGC-1α + XCT-790
H
relative mRNA levels
Atg9b
Ifrd1
shGFP + GFP
shGFP + PGC-1α
shErrα + PGC-1α + XCT-790

Figure 2. PGC-1α directly up-regulates both in a ERRα-dependent and -independent manner. (**A**) qRT-PCR analysis of PGC-1α, ERRα and Acadm mRNA levels in response to PGC-1α over-expression (OV) and shERRα knockdown (KD) + XCT-790. Data are normalized to mRNA levels in GFP infected cells. Error bars represent ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001. (**B**) Reverse cumulative distribution of log2 fold changes for all mouse promoters in the PGC-1α OV condition versus GFP control (left panel) and in the PGC-1α OV + shERRα KD + XCT-790 versus PGC-1α OV (right panel). Promoters are colored in red (up-regulation) when their fold change is bigger than 1.5 and in green (down-regulation) when their fold change is smaller than -1.5 (obtained by taking the inverse of the linear binding ratio). (**C**) Piechart representing the classification of directly up-regulated PGC-1α target genes in ERRα-dependent (orange) and ERRα-independent (yellow) targets. (**D**) Piechart representing the classification of indirectly up-regulated PGC-1α target genes in ERRα-dependent (violet) and ERRα-independent (lilac) targets. (**E-F**) Subset of the top significantly enriched GO terms identified for ERRα-dependent and ERRα-independent PGC-1α directly (**E**) or indirectly (**F**) induced target genes. (**G-H**) qRT-PCR analysis of two ERRα-dependent (**G**) or ERRα-independent (**H**) PGC-1α target genes, in response to PGC-1α OV and shERRα KD + XCT-790. Data are normalized to mRNA levels in GFP infected cells. Error bars represent ± SEM. *p < 0.05; **p < 0.01; ***p < 0.001.

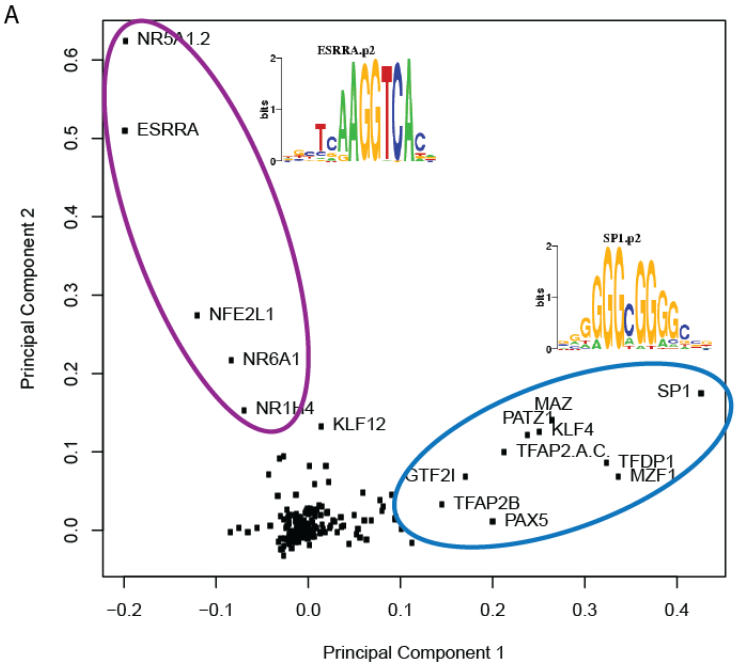
5.4.3. SP1 contributes to the up-regulation of ERRα targets in muscle cells

As a following step, we analyzed the occurrence of transcription factor DNA-binding motifs within all the ERRα peaks. We used the software MotEvo [14] to predict transcription factor binding sites (TFBSs) for a set of 190 known mammalian regulatory motifs downloaded from SwissRegulon [15]. In order to explain most of the binding site variation observed across the ERRα peaks, we applied principal component analysis (PCA) to a site-count matrix *N*, whose elements N_{pm} represent the number of predicted TFBSs for each motif *m* in each ERRα peak region *p*. Out of a total of 190, the first component was accounting for ~10% of the total variation in the dataset (Suppl. Fig. S2A). As shown in Figure 3A, the distribution of motif projections on the first two principal components clearly indicates two distinct clusters of motifs that stick out from the cloud of those having projections close to zero. The first group, characterized by motifs with negative eigenvalues along the first component, includes ESRRA and other nuclear receptors which have very similar logos to that of the ERRα motif; this cluster reflects the most abundant sites which can be found within the ERRα binding regions. Given the high promiscuity that exists between nuclear receptors consisting of different configurations of hexameric repeats [22], it might happen that the precise identity of the nuclear receptor bound at each peak is difficult to deduce from DNA-binding motifs and that, like in our case, a site for ERRα could be better matched by the position weight matrix (PWM) of another nuclear receptor. The second group, instead, is characterized by a set of GC-rich motifs which are usually found in the proximity of TSSs. The motif with the highest score along the first principal component corresponds to the binding affinity of the specificity protein 1 (SP1), known to be often present at the promoter region of many genes and to be involved in several cellular processes, including cell growth, cell differentiation, chromatin remodeling and apoptosis. The activity of this protein can be significantly affected by post-translational modifications, leading SP1 to act either as an

activator or as a repressor [23]. We thus asked whether SP1, in the context of ERRα binding, functions more as activator or as repressor on the common downstream target genes. For this purpose, we sought to combine the different classes of peaks (“only ERRα”, “only PGC-1α”, “overlapping ERRα and PGC-1α”, as defined in Fig. 1A) with the regulation of their assigned promoters (“up”, “down”, “non-changing”, “no promoter assigned”) in the plot shown in Suppl. Fig. S2B. Strikingly, whenever a site for SP1 is present within a peak, it is more likely for the assigned promoter to be up-regulated, strongly suggesting that in the context of PGC-1α over-expression, SP1 plays a role as an activator; this effect is enhanced when SP1 is found in an ERRα peak rather than in a PGC-1α peak.

An additional confirmation of the fact that the SP1 motif is particularly enriched in ERRα peaks is shown by Figure 3B, which displays SP1 as the top scoring motif when the TFBS occurrences in “only ERRα” peaks are compared to those predicted within “only PGC-1α” peaks. This observation is further supported by Figure 3C, showing the decrease in terms of TFBSs posterior sum for SP1, compared to the predictions in a shuffled peak set, when passing from “only ERRα” peaks to “overlapping ERRα and PGC-1α peaks” and finally to “only PGC-1α” peaks. Taken together, these results suggest not only a prevalent role of SP1 in the context of ERRα binding, but also that this co-occurrence at ERRα peaks is enriched when ERRα is not directly coactivated by PGC-1α, as SP1 occurs less often in the “only PGC-1α” peaks. Motivated by these results, we validated and confirmed by ChIP the presence of SP1 both at the promoters of the known target genes RIP140 and Fasn [24, 25] and at ERRα peaks hosting a predicted SP1 binding site in the proximity of four distinct protein coding genes (Fig. 3D). Another interesting observation from this binding sites motif search step is the difference in the number of SP1 TFBSs not only in the peaks, but also in the shuffled dataset (Fig. 3C); as SP1 is known to bind GC-rich regions, these results might reflect a different nucleotide composition between our three peak sets. Accordingly, we analyzed the CpG content of “only ERRα” and “only PGC-1α” peaks and further subdivided these peak sets into proximal and distal binding regions, where “proximal” is referred to peaks within 1kb from their associated gene promoter and “distal” to peaks located farther away. As clearly shown in Fig. 3E and Fig. 3F, the “only ERRα” peaks host more CpG dinucleotides with respect to “only PGC-1α” peaks; moreover, the fraction of “only ERRα” proximal peaks is much higher (~1/2) than the corresponding fraction of “only PGC-1α” peaks (~1/6). These results confirm the higher proximity of ERRα with respect to PGC-1α relative to their associated target genes and suggesting a different preference of ERRα binding to high- or low-CpG regions when recruited to proximal or distal sites, respectively. This pattern

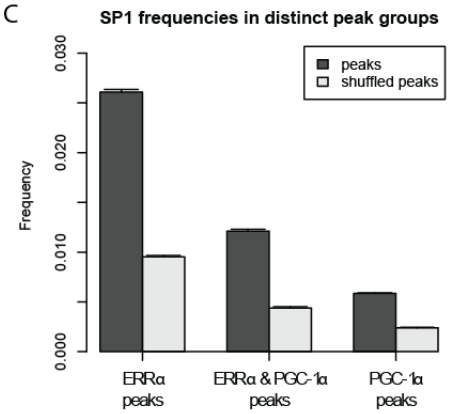
was reflected when we juxtaposed all ERRα peaks to all PGC-1α peaks by comparing their CpG- and GC-content (Suppl. Fig. S3).

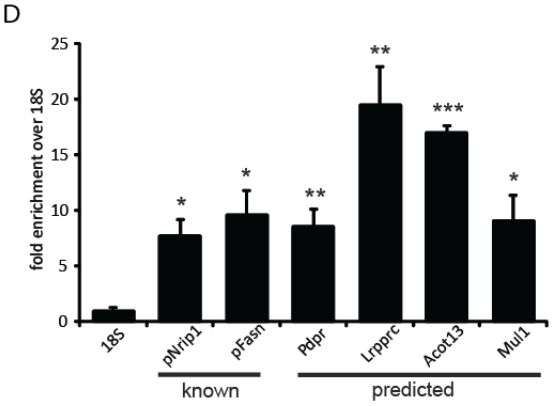


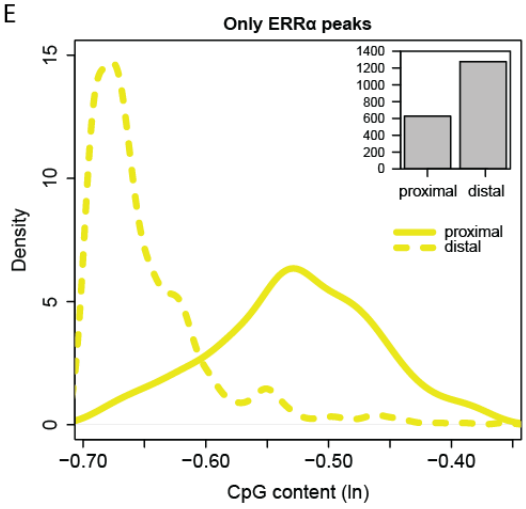
B

Only ERRα vs. Only PGC1α peaks

Motif name	Z score
SP1	75.743
PAX5	70.135
TFAP2B	65.739
HIC1	58.572
TFDP1	56.139
PATZ1	50.776
MZF1	50.112
KLF4	47.719
RREB1	46.779
TFAP2{A,C}	46.252
.	.
.	.
.	.
MEF2{A,B,C,D}	-30.566
JUN	-31.842
HOX{A6,A7,B6,B7}	-33.038
PAX8	-34.177
NR3C1	-34.908
AR	-36.214
NFE2	-37.495
NFE2L2	-40.473
FOSL2	-43.223
FOS_FOS{B,L1}_JUN{B,D}	-45.066
BACH2	-54.446







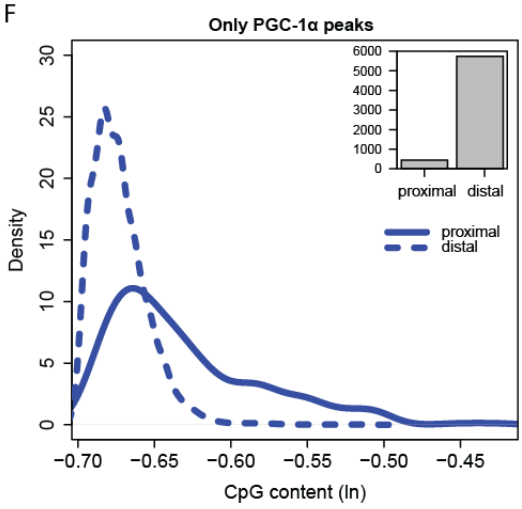


Figure 3. SP1 is the top transcription factor partner for ERRα in skeletal muscle. (**A**) PCA analysis of the 3225 ERRα peaks. The names of the motifs with the largest projections on the first two principal components are indicated. Purple and light blue ellipses highlight motif clusters, as identified by PC1, of nuclear hormone receptor-like motifs and SP1-like motifs, respectively. (**B**) Top and bottom scoring results of motif search obtained by comparing the TFBSs predictions within the "only ERRα peaks" with those in the "only PGC-1α peaks". (**C**) TFBSs posterior sum for SP1 in "only ERRα", "overlapping ERRα and PGC-1α" and "only PGC-1α" peaks. For each dataset, TFBS occurrences were compared against binding site predictions performed on the corresponding background set of shuffled peaks. (**D**) qRT-PCR validation of the ChIP enrichment measured at the promoter of a set of SP1 known target genes and around the predicted SP1 site within the ERRα peaks associated to the genes Pdpr, Lrpprc, Acot13 and Mul1. Bars represent fold enrichment over that of the 18S rRNA gene, error bars represent SEM. *$p < 0.05$; **$p < 0.01$; ***$p < 0.001$. (**E-F**) Density plot of the CpG content of "only ERRα" (**E**) and "only PGC-1α" (**F**) peaks, located either proximally (≤ 1 kb) or distally (> 1 kb) from the closest promoter. Each inset shows the bar plot of the number of "proximal" and "distal" peaks.

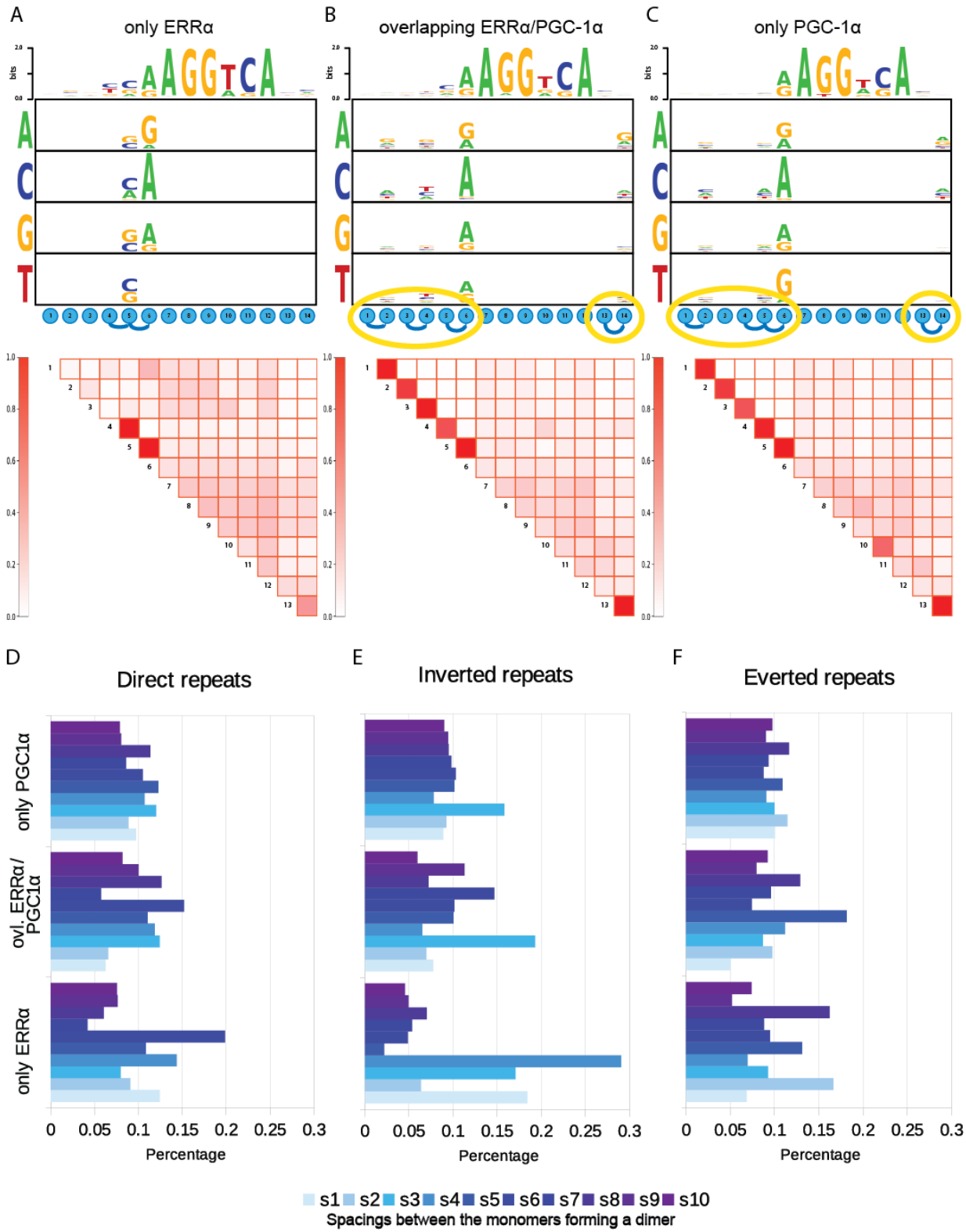
5.4.4. ERRα displays different binding specificities upon dimerization and PGC-1α coactivation

In the next step, we investigated whether ERRα showed distinct binding specificities across the three different peak groups and whether these might suggest precise preferences for a particular mode of binding (monomeric or dimeric) in a given genomic context. For this purpose, we used a novel approach for the analysis of transcription factor binding specificities, recently developed in our group (Omidi et al., in preparation). In particular, this method goes beyond the classical assumption that each nucleotide in a binding site contributes independently to the protein's interaction with DNA; instead, this approach takes into account the possible interdependencies between different positions, observed to be not negligible in several biochemical, statistical and crystallographic studies [26-28]. As shown by Omidi and colleagues, incorporating these dependencies into sequence-specific affinity models improves predictive accuracy in many cases.

Surprisingly, when we applied this approach to our three peak sets ("only ERRα", "overlapping ERRα and PGC-1α" and "only PGC-1α" peaks), we noticed that, although there are internal dependencies between the positions 4,5,and 6 in every group, the dependencies between initial and final positions (1-2 and 13-14) of the motif are only observed for "overlapping ERRα and PGC-1α" and "only PGC-1α" peaks, but not for "only ERRα" peaks (Fig. 4A-C). Since dependencies at the ends of the motif can be interpreted as an indication that the TF is likely to bind DNA as a dimer, these results suggest that, in the absence of direct coactivation by PGC-1α, ERRα binding sites might be found more often in a monomeric form rather than dimeric. We decided to go a step further and to investigate the presence of any difference in the spacing between potential dimers. For this purpose, we used the core recognition motif "AGGTCA" of the ESRRA weight matrix for building nuclear receptor dimers in direct, everted or inverted

configurations with a variable spacing between half-sites that ranged from 1 to 10 nucleotides. As shown in Figs.4 D-E, there was not a clearly preferred spacing for each particular dimer configuration; however, we noticed a general trend of some spacings more preferred than others in “only ERRα” and “overlapping ERRα and PGC-1α” peaks, as opposed to “only PGC-1α” peaks, suggesting the possibility that PGC-1α might coactivate several types of nuclear receptors with different spacing preferences between their monomers. In addition, the direct repeat dimer resulted as the most frequently observed configuration in all the three peak sets, predicted in 62.3% of the “only ERRα” peaks, in 58.9% of the “overlapping ERRα and PGC-1α” peaks and in 52.5% of the “only PGC-1α” peaks.

Interestingly, when we compared the fraction of monomers over dimers across peak sets (as shown in Fig. 4G), we noticed that this ratio was twice as big in “only ERRα” peaks with respect to the “overlapping ERRα and PGC-1α” peaks, suggesting that ERRα core recognition sites tend to occur more often as dimer when they are coactivated by PGC-1α. Furthermore, the same trend was observed when we normalized the number of monomers by the sum of monomers and dimers in each peak set; in fact, the “only ERRα” peaks showed the highest fraction of nuclear receptor monomers (39%) with respect to the other two groups (Fig. 4G). Taken together, our results suggest that the core recognition site of ERRα occurs more often in the monomeric form when it is not coactivated by PGC-1α, with respect to the case when both ERRα and PGC-1α co-localize in the same regulatory regions; moreover, these findings leave open the possibility that, as a monomer, ERRα might bind with other partners than PGC-1α.



G

	only ERRα peaks	ovl. ERRα/PGC-1α peaks	only PGC-1α peaks
posterior sum dimers	599.15	1354.69	3475.18
posterior sum monomers	378.88	432.60	1456.84
ratio monomers/dimers	0.63	0.32	0.42
ratio monomers/(monomers+dimers)	0.39	0.24	0.30

Figure 4. In the absence of a direct coactivation by PGC-1α, ERRα prefers to bind DNA as a monomer. (**A-C**) Motif logo showing the interdependencies between the different positions of the ESRRA weight matrix identified in "only ERRα", "overlapping ERRα and PGC-1α" and "only PGC-1α". Dependencies between positions are indicated by a blue curved line, while yellow ellipses highlight the dependencies which are in "overlapping ERRα and PGC-1α" and "only PGC-1α" peaks, but not in "only ERRα" peaks. (**D-F**) Percentage of direct, inverted and everted nuclear receptor dimers across our three peak sets, with spacings between half-sites ranging from 1 to 10. (**G**) Table showing the posterior sum and the fraction of nuclear receptor monomers and dimers across our three peak sets.

5.4.5. Modeling and validating the indirect gene regulatory effects of PGC-1α in absence of ERRα binding

In order to get an overview of the distinct peak distributions with regards to gene regulation, we stratified all differentially regulated target genes according to the types of ChIP-Seq peaks located in a window spanning from -10 kb from the gene's most upstream promoter to +10 kb from the gene's most downstream promoter. As shown in Figure 5A, the largest fraction of both up- and down-regulated genes does not have any PGC-1α and/or ERRα peaks in their neighborhood, suggesting that most of the differentially expressed genes are actually indirectly regulated by PGC-1α, in the absence of any ERRα peak. Therefore, we further investigated which other transcription factors might be mediating these indirect regulatory effects. For this reason, we extended a previously introduced framework, called Integrated System for Motif Activity Response Analysis (ISMARA) [19], to incorporate the ChIP-Seq data with the gene expression profiles (see Methods for more details) and to model these observations in terms of the TFBSs occurring within promoters and of the unknown regulatory "activities" of each input motif. In particular, after classifying promoters into distinct groups of direct or indirect targets, according to the presence or absence of PGC-1α and ERRα peaks, we focused particularly on the indirect targets group, whose top 30 motifs are listed in Figure 5B.

The first observation is that ESRRA is still the top motif of the list; this might be due to the above mentioned possibility of having homologous proteins recognized by the same PWM or to the fact that some real ERRα binding sites were just below the detection threshold of our peak caller. Secondly, we realized that the most active motifs host several known partners of either PGC-1α or ERRα: NR5A1 (also known as SF1), NR5A2 (also known as LHR1), MEF2C, RXRs and ESR1. Finally and most interestingly, among the top 10 motifs we noticed two transcription factors, Znf143 and Nfyc, which we previously suggested to be new potential PGC-1α partners and to be either significantly induced by PGC-1α or among the top over-represented motifs in the PGC-1α peaks [11]. In line with our previous findings, for both motifs the activity inferred by ISMARA increased significantly upon PGC-1α over-expression and decreased following ERRα knockdown in the subset of promoters not assigned to any peak, although to a lesser extent for

ZNF143 with respect to NFYC (Fig. 5C). Therefore, we decided to experimentally validate these two potential downstream effectors of PGC-1α by performing a siRNA-based knockdown of Znf143 and Nfyc in C2C12 myotubes and by analyzing, through a microarray, the expression levels of these two candidate transcription factors and of their target genes. This experiment resulted in a strong reduction of the expression of the TFs, as well as the expression levels of a total of 128 genes for Znf143 and 73 for Nfyc, which were all significantly induced by PGC-1α and repressed by the knockdown of the corresponding factor. Some examples of these Znf143 and Nfyc target genes, which were not assigned to any PGC-1α and/or ERRα peak, are shown in Fig. 5D. Furthermore, when we performed a Gene Ontology (GO) and KEGG over-representation analysis on the genes having no peaks around them, induced by PGC-1α and repressed by the knockdown of either Nfyc (81 genes) or Znf143 (154 genes), we noticed that the most enriched terms are related to mitochondrial organization and energy metabolism (Fig. 5E and 5F). These results highlight once more the central role of PGC-1α in the control of mitochondrial biogenesis and energy homeostasis in skeletal muscle, even in the absence of a direct recruitment of this coactivator. Moreover, since these gene sets represent ~15% and ~28% of Nfyc and Znf143 targets, respectively, we conclude that these two transcription factors are novel partners of PGC-1α in the indirectly mediated regulation of skeletal muscle plasticity.

A

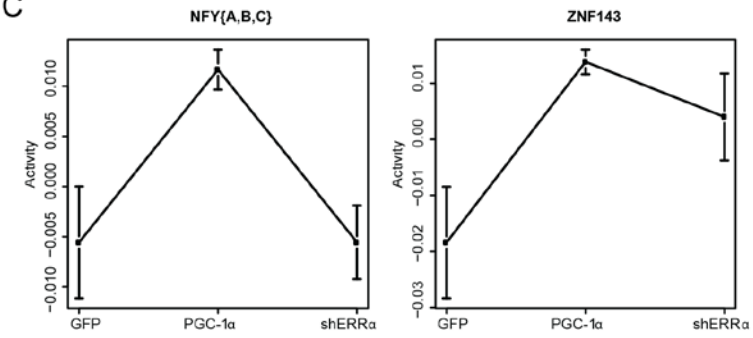
Type of peak	Up-regulated genes	Down-regulated genes
ERRα	164	6
PGC-1α	199	28
ERRα/PGC-1α	95	10
ERRα, PGC-1α	57	5
ERRα, ERRα/PGC-1α	16	0
PGC1α, ERRα/PGC-1α	57	1
ERRα, PGC-1α, ERRα/PGC-1α	21	2
none	556	417

B

ISMARA results in the context of neither ERRα nor PGC-1α peaks

Motif name	Z score
ESRRA.p2	9.012058
NR5A1,2.p2	6.217207
ESR1.p2	4.374093
TCF4_dimer.p2	4.016414
TFAP2B.p2	3.964363
CTCF.p2	3.961013
ZNF143.p2	**3.772258**
SRF.p3	3.649944
NFY{A,B,C}.p2	**3.522142**
TLX1..3_NFIC{dimer}.p2	3.486428
ZNF384.p2	3.442566
TFCP2.p2	3.40105
RREB1.p2	3.340566
RXR{A,B,G}.p2	3.332749
E2F1..5.p2	3.305118
NFATC1..3.p2	3.182886
TBP.p2	3.122774
STAT2,4,6.p2	3.040521
POU5F1.p2	3.021371
ZIC1..3.p2	2.978738
GATA1..3.p2	2.97715
IRF1,2,7.p3	2.826192
SOX2.p2	2.805607
SPIB.p2	2.746749
PRDM1.p3	2.728316
ATF6.p2	2.625964
FOXD3.p2	2.573846
FOX{I1,J2}.p2	2.5151
MEF2{A,B,C,D}.p2	2.509546
ZFP161.p2	2.473862

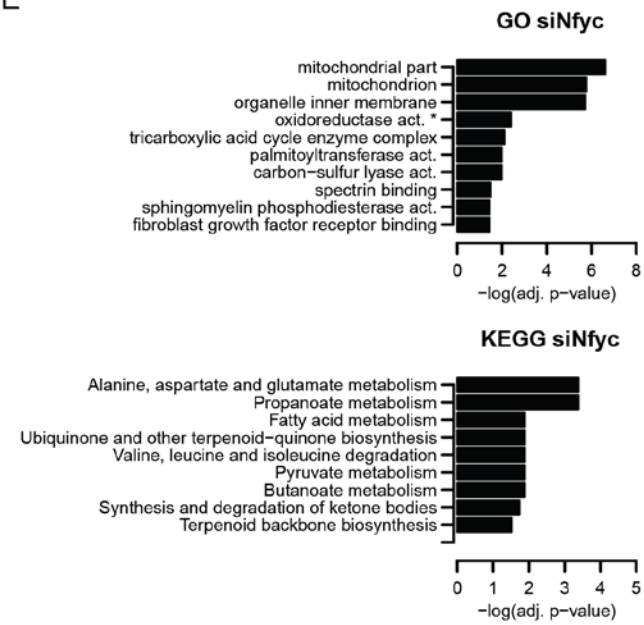
C



D

	PGC-1α OV	siNfyc KD		PGC-1α OV	siZnf143 KD
Nfyc	1.89	-2.60	Znf143	1.28	-1.73
Ifrd1	2.84	-1.8	Tle4	2.08	-2.98
Hmgb3	2.03	-1.65	Acadvl	3.42	-2.73
Ddo	2.83	-3.53	Gyk	4.29	-1.92
Eaf2	2.7	-1.55	Mrpl37	2.15	-1.48
Gas2l3	4.29	-1.66	Gm5113	4.18	-2.07

E



F

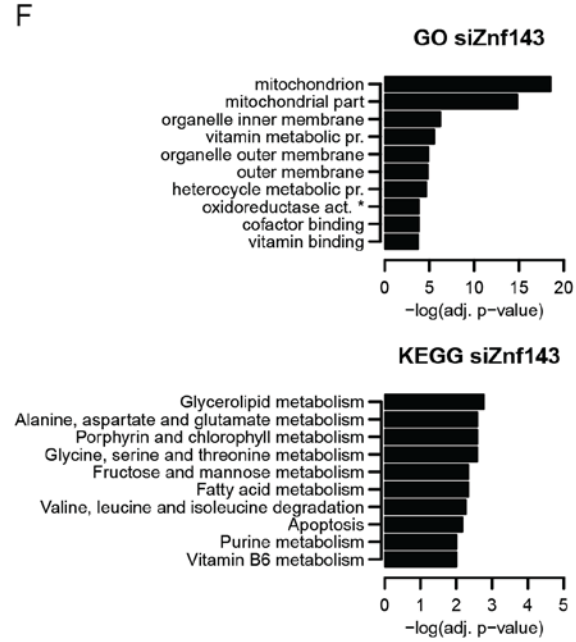


Figure 5. Prediction and validation of the indirect gene regulatory effects of PGC-1α, in the absence of ERRα recruitment, mediated by Nfyc and Znf143. (**A**) Counts of different combinations of PGC-1α and ERRα peaks located in the vicinity of all differentially regulated genes. (**B**) Top 30 active motifs in the promoters not associated to any PGC-1α and/or ERRα peak, as predicted by ISMARA. The motifs chosen for validation are highlighted in bold. (**C**) Activity plots of the motifs NFY(A,B,C) and ZNF143, as predicted by ISMARA, in the following conditions: (i) GFP over-expression, (ii) PGC-1α over-expression and (iii) PGC-1α over-expression + shERRα knockdown + XCT-790. (**D**) Fold change levels, as measured by a microarray experiment, of Nfyc, Znf143 and 10 other genes significantly up-regulated by PGC-1α over-expression (OV), significantly down-regulated by either siNfyc or siZnf143 knockdown (KD) and without any PGC-1α and/or ERRα peak in their vicinity. (**E-F**) Subsets of the top significantly enriched GO and KEGG terms identified for Nfyc (**E**) and Znf143 (**F**) target gene sets, as defined in Fig.5D.

5.5. Discussion

The nuclear receptor ERRα has been described as an effector of PGC-1α-induced regulation of gene expression in all highly oxidative organs, like skeletal muscle, adipose tissue or heart [29-31]. In skeletal muscle, PGC-1α promotes and coactivates ERRα to increase its transcriptional activity; in turn, these two proteins were shown to induce the expression of genes involved in mitochondrial biogenesis, oxidative phosphorylation and fatty acid oxidation [7, 8, 32]. Moreover, the induction of the vascular endothelial growth factor (VEGF) and angiogenesis in skeletal muscle by PGC-1α involves transcriptional activity of ERRα [33]; in addition, PGC-1α induces the sirtuin 3 (SIRT3) gene through ERRα to regulate the suppression of reactive oxygen species in skeletal muscle cells [34].

However, all these studies were focused on the transcription of individual genes or of a small set of genes. To date, it is not known, on a genome-wide scale, to which extent is ERRα required by PGC-1α for the regulation of these biological programs. We now answered this question by providing a global picture of gene expression regulation by PGC-1α and ERRα in skeletal muscle cells. The first evidence that ERRα can be recruited to DNA also independently from PGC-1α is provided by the fact that our ChIP-Seq experiment revealed binding of ERRα and PGC-1α to both shared and distinct sets of target genes. Indeed, while the occurrence of ERRα at the promoter regions of known PGC-1α target genes (Suppl. Fig. S1) confirms that PGC-1α coactivates ERRα to drive the expression of these genes, the observation that only a minor fraction of ERRα and PGC-1α peaks is overlapping suggests that the long-term believed concept of a symbiotic cooperation between these two proteins is actually only restricted to a subset of their target genes. Moreover, we showed that ERRα is required by PGC-1α to induce the expression of about two thirds of both its directly- and indirectly-controlled targets. These results confirm that ERRα is a crucial partner of PGC-1α and that it is required for the regulation

of most of PGC-1α targets, in order to guarantee an efficient control of processes which are essential for cell energy metabolism. On the other hand, PGC-1α is able to control these biological pathways also in the absence of its classical binding partner and, probably, through coactivation of alternative transcription factors. It is in fact not surprising that PGC-1α can be recruited to the regulatory elements of its target genes without being bound to ERRα, as several studies previously showed that PGC-1α is capable of binding and coactivating numerous transcription factors in skeletal muscle. Some examples are the nuclear respiratory factor 1 (NRF-1) and the nuclear respiratory factor 2 (NRF-2). Upon coactivation by PGC-1α, the former TF induces the expression of the mitochondrial transcription factor A (mtTFA) and thus the transcription and replication of mitochondrial DNA [35], while the latter TF promotes the expression of genes involved in oxidative phosphorylation and neuromuscular junction formation [8, 36]. Similarly, PGC-1α regulates the activity of the myocyte enhancer factor 2C (MEF2C) to induce the expression of GLUT4 and thus to control glucose uptake in muscle [37]. Nevertheless, although it was expected that PGC-1α can regulate transcription also without ERRα, we were surprised to find that, symmetrically, ERRα can be recruited to DNA also without direct coactivation by PGC-1α. Previous studies, showing that ERRα is in a transcriptionally active conformation when complexed with PGC-1α in the absence of other ligands [38], led to the idea that this orphan nuclear receptor is activated by coactivator proteins rather than by small molecule ligands [39]. Here we show, on a genome-wide scale, that ERRα can bind to several regulatory regions without a direct coactivation by PGC-1α.

A reasonable question at this point would be: if ERRα is often not directly coactivated by PGC-1α, can we identify other possible partners of ERRα co-occurring at the same regulatory regions? By combining motif search and principal component analysis (PCA), we showed here that SP1 is the top over-represented transcription factor in the subset of ERRα peaks not overlapping any PGC-1α peak. In addition, we validated by ChIP the significant enrichment of SP1 at four predicted SP1 binding sites within ERRα peaks. To date, there is only one evidence that SP1 can interact in vivo with ERRα, as shown by a GST pull-down assay in 2006 [40], although it is known that the GST protein is prone to non-specific interactions [41]. Finally, another study identified both SP1 and ERRα as modulators of the transcriptional activity of the coregulator RIP140, also called Nrip1 [24]. Herewith, we provide clear evidence, by ChIP validation, that SP1 and ERRα are co-localized at the same regulatory regions.

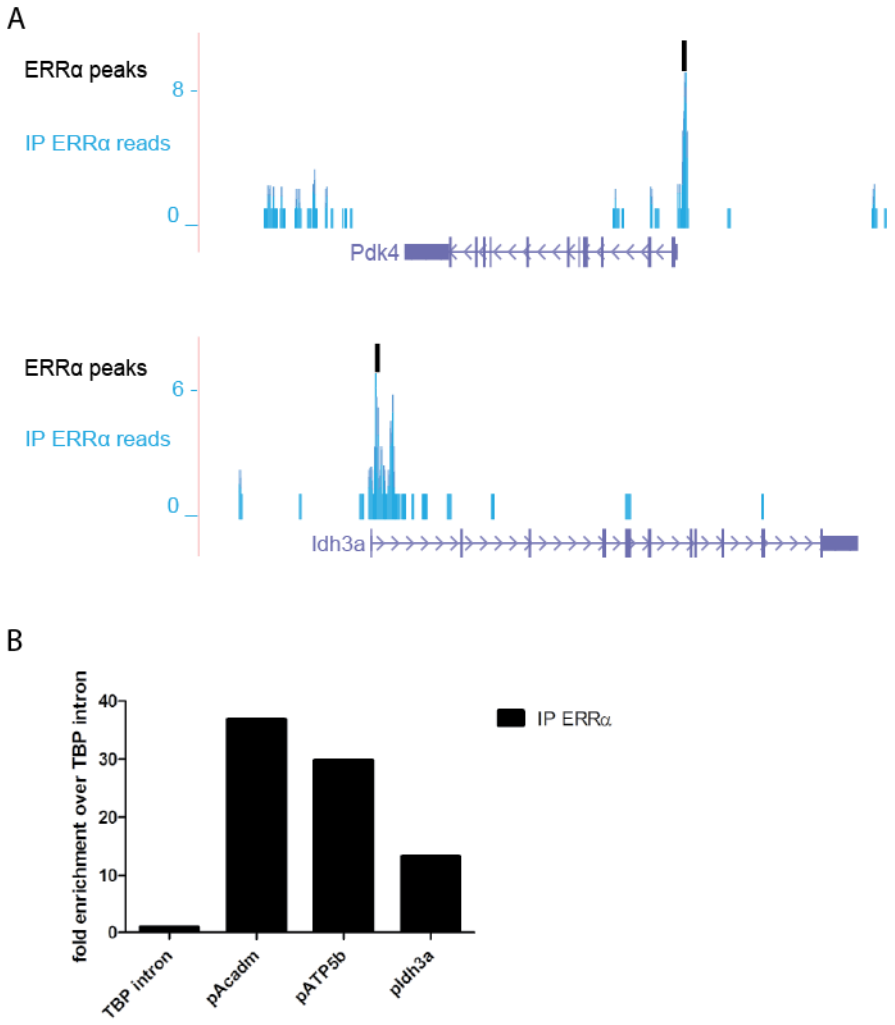
Most interestingly, we also found that SP1 tends to occur significantly more often in “only ERRα” peaks than in “only PGC-1α” peaks. In addition, we observed that the binding affinity of ERRα itself is slightly different in these distinct peak sets and that, in “only ERRα” peaks, nuclear

receptors (including ERRα) tend to occur more often as monomers rather than as dimers. These findings, which require further investigation, would suggest that ERRα might function in a dual way: in GC-poor regions, it is often binding DNA as a dimer and it is directly coactivated by PGC-1α; on the other hand, in GC-rich regions, ERRα mostly co-occurs as a monomer together with SP1. A previous study by Barry and coworkers showed that minor changes in the ERRα response element (ERRE) are sufficient to determine a different preference in the mode of binding (as monomer or dimer) and therefore to affect ERRα interaction with the coactivator PGC-1α [42]. Our study now strengthens this concept with DNA occupancy data and extends it to the skeletal muscle tissue on a genome-wide scale.

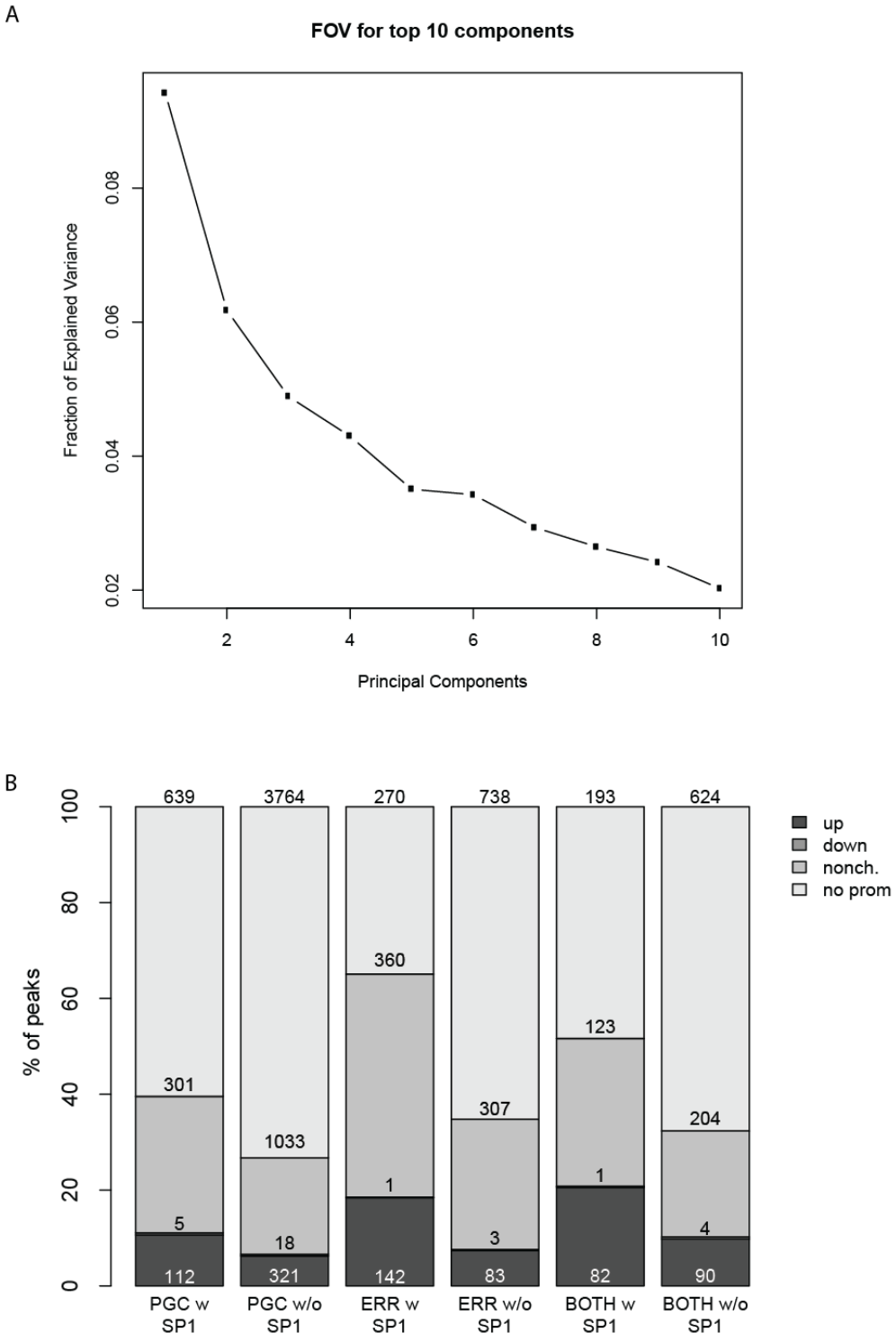
Finally, we investigated which other transcription factors might mediate the regulation of indirect PGC-1α targets in skeletal muscle without the recruitment of ERRα to their promoters. These indirect PGC-1α target genes might be regulated by transcription factors or coactivators which are themselves induced, but not coactivated by PGC-1α. One such downstream effector of the PGC-1α/ERRα signaling axis is Perm1, which is induced by PGC-1α and subsequently regulates the expression of some PGC-1α target genes [30]. Our results now indicate that the transcription factors Znf143 and Nfyc might be other downstream effectors of PGC-1α. In a previous study [11], we showed that the knockdown of these two TFs was able to prevent the PGC-1α-controlled up-regulation of a set of genes associated to PGC-1α peaks with binding site predictions for either Znf143 or Nfyc. Now, we show, by an expression array experiment, that two other subsets of indirect PGC-1α target genes – that is, not assigned to any PGC-1α peak – are significantly down-regulated by either Znf143 or Nfyc knockdown, respectively.

In conclusion, this study describes the genome-wide map of ERRα recruitment to DNA, in the context of over-expressed PGC-1α in skeletal muscle. By a comparison to a previous study on PGC-1α occupancy, we could identify and distinguish different subsets of genes whose regulatory regions were either bound by ERRα, PGC-1α or both. Moreover, we discovered that ERRα is required by PGC-1α to induce most of its direct and indirect targets, mainly involved in the regulation of mitochondrial organization and energy metabolism. In addition, we propose a mechanism in which ERRα is coactivated mostly as a dimer by PGC-1α, whereas, as a monomer, it is often co-localized with SP1 in GC-rich regions. Finally, we suggest two transcription factors, Nfyc and Znf143, as novel potential mediators of PGC-1α indirectly regulated gene up-regulation. In summary, by elucidating the relationship between PGC-1α and ERRα, we additionally identified novel members and regulatory mechanisms of the PGC-1α-controlled transcriptional network that modulates energy metabolism in skeletal muscle.

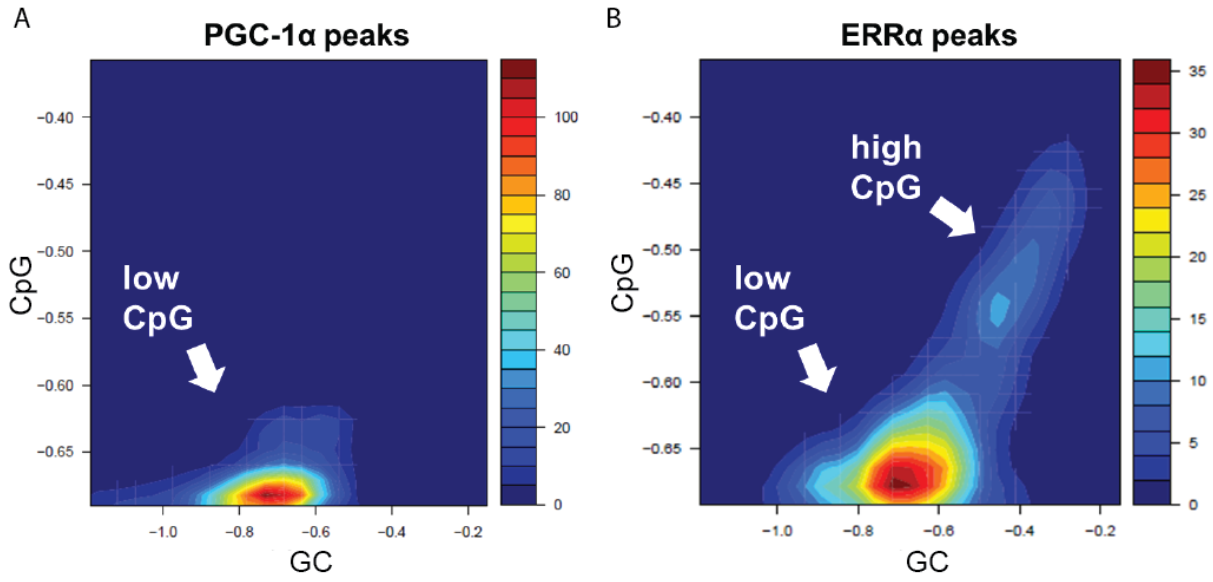
5.6. Supplementary figures and tables



Suppl. Figure S1. Related to Figure 1. (**A**) ERRα read densities around the TSS of the known target genes Pdk4 and Idh3a, as displayed by the UCSC Genome Browser. (**B**) Real-time PCR validation of the ChIP enrichment measured at the promoter of a set of ERRα target genes. Bars represent fold enrichment over that of the TBP intron.



Suppl. Figure S2. Related to Figure 3. (**A**) Fraction of explained variance of the top 10 PCA components. (**B**) Barchart representing the different classes of peaks ("only ERRα", "only PGC1α", "overlapping ERRα and PGC1α") together with the regulation of their associated promoters ("up", "down", "non-changing", "no promoter assigned"). Numbers shown on top of each box represent the absolute peak counts.



Suppl. Figure S3. Related to Figure 3. (**A-B**) Two-dimensional histogram (shown as a heatmap) of the GC base content (horizontal axis) and CpG dinucleotide content (vertical axis) of all PGC-1α peaks (**A**) and of all ERRα peaks (**B**). The values shown on both axes are expressed as logarithms.

Suppl. Table 1. Real-time primer sequences

Real-time PCR primers used for ChIP validation		
Gene promoter or intron	*Forward primer*	*Reverse primer*
TBP intron	TGTGAGCTCCTTGGCTTTTT	ATAGTTGCCCAGCAATCAGG
promoter of Acadm	CCTTGCCCGAGCCTAAAC	GTCTGGCTGCGCCCTCT
promoter of ATP5b	CTGGAAACTTCCACCCTCACTA	GAGAGGTTTTTGGCGGAACTA
promoter of Idh3a	GGACGGCGTCAAGGTCAAG	GCCTAGGTGGCCTGTCTGTG
pNrip1	CACGCCATTCAGCTCTTCAG	GTGACAATGGGAGGGAGGG
pFasn	CTGGAGCACAAGGAACGC	GGACAGAGATGAGGGCGTC
Pdpr	CACACTCGTCGTCAACCAG	GTGCGCTTGTTTGGGTCTC
Lrpprc	ACAACACCCCTCCACTTTGA	CGGTGTCGCTCCTAGTTG
Acot13	TCACTCTTTAGCGCCCCAG	AAGACCGCCCTCTCTGGT
Mul1	ACTCCATATACCGGCAGAAGG	GAGCTGCCAGTGAGACCG

Real-time PCR primers used for testing the knockdown of ERRα		
Gene	*Forward primer*	*Reverse primer*
18S	AGTCCCTGCCCTTTGTACACA	CGATCCGAGGGCCTCACTA
PGC-1α	TGATGTGAATGACTTGGATACAGACA	GCTCATTGTTGTACTGGTTGGATATG
ERRα	ACTGCAGAGTGTGTGGATGG	GCCCCCTCTTCATCTAGGAC
Acadm	AACACTTACTATGCCTCGATTGCA	CCATAGCCTCCGAAAATCTGAA
Aim1l	CCTGTTGCGTCCATAAGGGT	GCTCTGAGTTCCACATCCCC
Twf2	TGCTACCTCCTCTTCCGACT	ATAGCATCTTCAGCCGCACC
Atg9b	TGGCATCACATCCAGAACCT	CATTGTAATCCACGCAGCGA
Ifrd1	GACAAGAGAAAGCAGCGGTC	GGTACTGCATCCCTGATCCA

References

1. Puigserver, P. and B.M. Spiegelman, *Peroxisome proliferator-activated receptor-gamma coactivator 1 alpha (PGC-1 alpha): transcriptional coactivator and metabolic regulator.* Endocrine reviews, 2003. 24(1): p. 78-90.
2. Handschin, C., *Regulation of skeletal muscle cell plasticity by the peroxisome proliferator-activated receptor gamma coactivator 1alpha.* Journal of receptor and signal transduction research, 2010. 30(6): p. 376-84.
3. Handschin, C. and B.M. Spiegelman, *Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism.* Endocrine reviews, 2006. 27(7): p. 728-35.
4. Schreiber, S.N., et al., *The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor estrogen-related receptor alpha (ERRalpha).* The Journal of biological chemistry, 2003. 278(11): p. 9013-8.
5. Wende, A.R., et al., *PGC-1alpha coactivates PDK4 gene expression via the orphan nuclear receptor ERRalpha: a mechanism for transcriptional control of muscle glucose metabolism.* Molecular and cellular biology, 2005. 25(24): p. 10684-94.
6. Liu, D., H. Benlhabib, and C.R. Mendelson, *cAMP enhances estrogen-related receptor alpha (ERRalpha) transcriptional activity at the SP-A promoter by increasing its interaction with protein kinase A and steroid receptor coactivator 2 (SRC-2).* Molecular endocrinology, 2009. 23(6): p. 772-83.
7. Schreiber, S.N., et al., *The estrogen-related receptor alpha (ERRalpha) functions in PPARgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6472-7.
8. Mootha, V.K., et al., *Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6570-5.
9. Zhang, Z. and C.T. Teng, *Estrogen receptor-related receptor alpha 1 interacts with coactivator and constitutively activates the estrogen response elements of the human lactoferrin gene.* The Journal of biological chemistry, 2000. 275(27): p. 20837-46.
10. Kamei, Y., et al., *PPARgamma coactivator 1beta/ERR ligand 1 is an ERR protein ligand, whose expression induces a high-energy expenditure and antagonizes obesity.* Proceedings of the National Academy of Sciences of the United States of America, 2003. 100(21): p. 12378-83.
11. Baresic, M., et al., *Transcriptional network analysis in muscle reveals AP-1 as a partner of PGC-1α in the hypoxic gene program (manuscript currently in press)*, 2013.
12. Langmead, B., et al., *Ultrafast and memory-efficient alignment of short DNA sequences to the human genome.* Genome biology, 2009. 10(3): p. R25.
13. Notredame, C., D.G. Higgins, and J. Heringa, *T-Coffee: A novel method for fast and accurate multiple sequence alignment.* Journal of molecular biology, 2000. 302(1): p. 205-17.
14. Arnold, P., et al., *MotEvo: integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences.* Bioinformatics, 2012. 28(4): p. 487-94.
15. Pachkov, M., et al., *SwissRegulon, a database of genome-wide annotations of regulatory sites: recent updates.* Nucleic acids research, 2013. 41(Database issue): p. D214-20.
16. R Development Core Team, *R: A language and environment for statistical computing.*, 2012, R Foundation for Statistical Computing.

17. Gentleman, R.C., et al., *Bioconductor: open software development for computational biology and bioinformatics.* Genome biology, 2004. 5(10): p. R80.
18. Al-Shahrour, F., R. Diaz-Uriarte, and J. Dopazo, *FatiGO: a web tool for finding significant associations of Gene Ontology terms with groups of genes.* Bioinformatics, 2004. 20(4): p. 578-80.
19. Suzuki, H., et al., *The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line.* Nature genetics, 2009. 41(5): p. 553-62.
20. Tiraby, C., et al., *Estrogen-related receptor gamma promotes mesenchymal-to-epithelial transition and suppresses breast tumor growth.* Cancer research, 2011. 71(7): p. 2518-28.
21. Zhang, Y., et al., *Estrogen-related receptors stimulate pyruvate dehydrogenase kinase isoform 4 gene expression.* The Journal of biological chemistry, 2006. 281(52): p. 39897-906.
22. Mangelsdorf, D.J. and R.M. Evans, *The RXR heterodimers and orphan receptors.* Cell, 1995. 83(6): p. 841-50.
23. Song, C.Z., et al., *Functional interaction between coactivators CBP/p300, PCAF, and transcription factor FKLF2.* The Journal of biological chemistry, 2002. 277(9): p. 7029-36.
24. Nichol, D., et al., *RIP140 expression is stimulated by estrogen-related receptor alpha during adipogenesis.* The Journal of biological chemistry, 2006. 281(43): p. 32140-7.
25. Samson, S.L. and N.C. Wong, *Role of Sp1 in insulin regulation of gene expression.* Journal of molecular endocrinology, 2002. 29(3): p. 265-79.
26. Berger, M.F., et al., *Compact, universal DNA microarrays to comprehensively determine transcription-factor binding site specificities.* Nature biotechnology, 2006. 24(11): p. 1429-35.
27. Luscombe, N.M., R.A. Laskowski, and J.M. Thornton, *Amino acid-base interactions: a three-dimensional analysis of protein-DNA interactions at an atomic level.* Nucleic acids research, 2001. 29(13): p. 2860-74.
28. Zhou, Q. and J.S. Liu, *Modeling within-motif dependence for transcription factor binding site predictions.* Bioinformatics, 2004. 20(6): p. 909-16.
29. Akter, M.H., et al., *Perilipin, a critical regulator of fat storage and breakdown, is a target gene of estrogen receptor-related receptor alpha.* Biochemical and biophysical research communications, 2008. 368(3): p. 563-8.
30. Cho, Y., et al., *Peroxisome Proliferator-activated Receptor gamma Coactivator 1 (PGC-1)- and Estrogen-related Receptor (ERR)-induced Regulator in Muscle 1 (PERM1) Is a Tissue-specific Regulator of Oxidative Capacity in Skeletal Muscle Cells.* The Journal of biological chemistry, 2013. 288(35): p. 25207-18.
31. Huss, J.M., R.P. Kopp, and D.P. Kelly, *Peroxisome proliferator-activated receptor coactivator-1alpha (PGC-1alpha) coactivates the cardiac-enriched nuclear receptors estrogen-related receptor-alpha and -gamma. Identification of novel leucine-rich interaction motif within PGC-1alpha.* The Journal of biological chemistry, 2002. 277(43): p. 40265-74.
32. Huss, J.M., et al., *Estrogen-related receptor alpha directs peroxisome proliferator-activated receptor alpha signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle.* Molecular and cellular biology, 2004. 24(20): p. 9079-91.
33. Arany, Z., et al., *HIF-independent regulation of VEGF and angiogenesis by the transcriptional coactivator PGC-1alpha.* Nature, 2008. 451(7181): p. 1008-12.
34. Kong, X., et al., *Sirtuin 3, a new target of PGC-1alpha, plays an important role in the suppression of ROS and mitochondrial biogenesis.* PloS one, 2010. 5(7): p. e11707.
35. Wu, Z., et al., *Mechanisms controlling mitochondrial biogenesis and respiration through the thermogenic coactivator PGC-1.* Cell, 1999. 98(1): p. 115-24.

36. Handschin, C., et al., *PGC-1alpha regulates the neuromuscular junction program and ameliorates Duchenne muscular dystrophy.* Genes & development, 2007. 21(7): p. 770-83.
37. Michael, L.F., et al., *Restoration of insulin-sensitive glucose transporter (GLUT4) gene expression in muscle cells by the transcriptional coactivator PGC-1.* Proceedings of the National Academy of Sciences of the United States of America, 2001. 98(7): p. 3820-5.
38. Kallen, J., et al., *Evidence for ligand-independent transcriptional activation of the human estrogen-related receptor alpha (ERRalpha): crystal structure of ERRalpha ligand binding domain in complex with peroxisome proliferator-activated receptor coactivator-1alpha.* The Journal of biological chemistry, 2004. 279(47): p. 49330-7.
39. Gaillard, S., et al., *Receptor-selective coactivators as tools to define the biology of specific receptor-coactivator pairs.* Molecular cell, 2006. 24(5): p. 797-803.
40. Castet, A., et al., *Receptor-interacting protein 140 differentially regulates estrogen receptor-related receptor transactivation depending on target genes.* Molecular endocrinology, 2006. 20(5): p. 1035-47.
41. Walker, J.M. and R. Rapley, *Medical biomethods handbook.* 2005, Totowa, N.J.: Humana Press. xv, 644 p.
42. Barry, J.B., J. Laganiere, and V. Giguere, *A single nucleotide in an estrogen-related receptor alpha site can dictate mode of binding and peroxisome proliferator-activated receptor gamma coactivator 1alpha activation of target promoters.* Molecular endocrinology, 2006. 20(2): p. 302-10.

6. Discussion

More than a decade of studies have established the coactivator PGC-1α and its well-known partner ERRα as the key modulators of skeletal muscle plasticity in response to a myriad of factors – including exercise, caloric restriction and oxidative stress – and with major consequences for health (for a recent review, see [1]). To address such a complex variety of stimuli, these proteins interact with a multitude of chromatin remodelers and transcription factors, of which, most likely, only a few have been discovered so far. In this study, we expand the transcriptional network of PGC-1α and ERRα through a successful combination of high-throughput techniques and computational analysis. Using a genome-wide approach, for the first time to our knowledge, we have mapped all the loci in skeletal muscle cells where PGC-1α and ERRα are recruited to. Moreover, the fruitful integration of ChIP-Seq and microarray data allowed us to link genomic occupancy to direct and indirect target gene regulation.

The first focus of this study was the optimization of the several steps involved in our ChIP-Seq analysis pipeline. In particular, we worked out a series of artifacts one could encounter in this type of computational analysis. One of these is the PCR amplification error, caused by extra runs of amplification from only few starting DNA molecules, which may lead to problems in the downstream steps of the analysis. Another such artifact originates from the under-estimated presence of repeats in reference genomes, causing lots of sequenced reads to map all in the same region and giving rise to false positive peaks. Moreover, all current high-throughput technologies have an average error rate that is considerably higher than the classical Sanger sequencing [2], requiring a quality filtering step before the read mapping to the reference genome. In our opinion, the best way to deal with these problems was, respectively: (i) to assign a weight of 1 to any “stack” of reads starting at the same genomic position, (ii) to remove sliding windows containing an amount of background reads exceeding a given cutoff and (iii) to implement a read quality filter based on the FASTQ sequencing scores. In addition, to reduce the amount of false positive peaks called by commonly used peak finders, we implemented our own algorithm to evaluate local IP enrichment versus background by calculating a Z score for every genomic region. In the future developments of this study, we could use a more sophisticated noise model for the peak calling step (for example, a log-normal distribution of read counts) and adopt a mixture model to refine peaks and to identify individual binding events at higher resolution.

One of the our first observations from the analysis of PGC-1α and ERRα ChIP-Seq experiments was the fact that both these proteins can often be found in intergenic regions, which is much further from genes than what current literature suggests; in fact, most of the genetic and molecular studies on PGC-1α and ERRα identified their recruitment sites in the immediate vicinity (or proximal promoters) of their target genes (for example, see [3, 4]). Most interestingly, the majority (~57%) of our ERRα peaks – and an even higher fraction of PGC-1α peaks (~69%) – is not assigned to any mouse promoter within a distance of 10 kb. Because of this maximum-distance cutoff, one limitation of our method is to consider the peak as not annotated whenever it is located further than this distance from the closest promoter; however, a considerable fraction of our peaks (~25% for ERRα and ~31% for PGC-1α) are located within an intronic region of a very long gene, thus being more distant than the defined cutoff from the gene's promoter. As shown by several studies (for example [5]), *cis*-acting regulatory elements located within noncoding regions of genomic DNA, like introns, can loop over DNA and enable long-range regulatory interactions with promoter regions. In addition, we showed that a considerable fraction of these not annotated peaks overlap histone modifications like H3K4me1 and H3K18ac, both known to be present at active and inactive/poised enhancers (see chapter 3). Taken together, these considerations suggest the hypothesis that several of our PGC-1α and ERRα peaks might be located within intergenic or intronic enhancers and thus behave as distal regulatory elements.

Another interesting outcome of our study is the confirmation of the inability of PGC-1α to directly repress its downstream targets expression, in contrast to what was suggested by other studies [6, 7]. Using a customized version of ISMARA [8], we integrated the information coming from an expression array of over-expressed PGC-1α versus GFP to the ChIP-Seq data; in this way, we were able to model, for the first time, the direct and indirect regulatory effects of PGC-1α in skeletal muscle. Although several motifs showed a significant decrease of their activity upon PGC-1α over-expression on indirect targets, no significant reduction in the activity exerted on direct target promoters was observed for any motif. Most interestingly, the group of transcription factors showing this activity pattern includes the STAT proteins, the IRFs and NFκB, all known to be involved in inflammatory processes. Recently, we and others have shown that PGC-1α is able to reduce the transcriptional activity of NFκB in vitro and in vivo [9, 10]. Our results now confirm that this anti-inflammatory activity of PGC-1α is exclusively indirect. Moreover, as shown in chapter 3, the fact that the overlap between PGC-1α non-annotated peaks and the repressive histone modification H3K27me3 was one of the lowest we observed (only ~4.9%), suggests that an alternative repressive mechanism of modifying chromatin condensation from an active to a repressive state can also be excluded.

By distinguishing direct from indirect effects of PGC-1α, we also found, unexpectedly, that many genes involved in the highly enriched GO terms related to oxidative metabolic pathways are indirectly controlled by PGC-1α, meaning that they do not require a direct recruitment of this coactivator to their regulatory sites. Based on previous studies, we expected the regulation of these pathways to be directly controlled by PGC-1α, mostly in partnership with ERRα [11, 12]. Surprisingly, our data indicate that ERRα is active both in presence and absence of PGC-1α coactivation (see chapter 4), suggesting that the nuclear receptor was already induced, most probably by PGC-1α itself, at the one time point when the experiment was performed. Altogether, these observations suggest that PGC-1α might affect oxidative metabolic pathways at different levels, primarily acting as an upstream regulator of other TFs (like ERRα), which are controlling, at later stages, more downstream PGC-1α targets without need of a direct coactivation by PGC-1α itself. Unfortunately, a ChIP-Seq experiment is merely a snapshot of the occupancy of a protein of interest at a given time point. On the other hand, taking into account the fact that PGC-1α is rhythmically expressed in skeletal muscle and stimulates the expression of clock genes like Bmal1 and Rev-erbα [13], an interesting future direction of this project would be the analysis of PGC-1α occupancy in skeletal muscle at different time points. This way, integrating these data to a time-course RNA profiling dataset like the one described in [14], we could identify the temporal and causal relationship between PGC-1α and all its intermediate effectors that lead to the activation of mitochondrial biogenesis and energy metabolism in skeletal muscle.

Regarding ERRα, several studies have focused so far on the regulation of individual genes or small sets of genes. Our study provides, for the first time to our knowledge, a global picture of ERRα genome-wide recruitment in skeletal muscle cells, in the context of PGC-1α over-expression (see chapter 5). As mentioned before, ERRα seems to be able to bind and regulate target gene expression also without being directly coactivated by PGC-1α. Although previous studies showed that ERRα is induced to high levels by PGC-1α and is present in an active conformation when bound to PGC-1α [15, 16], there are no evidences to date that this nuclear receptor can also be recruited to the DNA without a direct interaction with PGC-1α. In a report from Ronald Evans' group, the authors demonstrated that the transcriptional activity of hERR1, the human ortholog of the murine ERRα, is potentiated in a ligand-independent manner by the coactivators ACTR (activator of thyroid and retinoic acid receptors), GRIP1 (glucocorticoid receptor interacting protein 1), and SRC-1 (steroid receptor coactivator 1) [17]. A later study in human breast cancer showed that the correlation between ERRα expression with that of the coactivator ACTR (also called AIB1) is even stronger than with PGC-1α [18]. Taken together,

these studies strongly suggest that, apart from PGC-1α, other coactivators could as well regulate ERRα transcriptional activity in skeletal muscle, therefore explaining the lack of PGC-1α recruitment to most of the ERRα sites.

An obvious question at this point is the following: if PGC-1α is not coactivating at these ERRα peaks, what other factors is ERRα working with? As shown in chapter 5, the specificity factor SP1 turned out to be the most over-represented transcription factor within the peaks occupied by ERRα, but not by PGC-1α. Moreover, we could validate some of our SP1 binding site predictions and show a significant enrichment of this protein at some ERRα peaks. The few studies that have previously investigated the relationship between ERRα and SP1 are somehow contradictory. The first, in 2003, proposed that SP1 is important for ERRα transcription in human, although the authors did not show actual binding of the SP1 protein to the corresponding sites at the ERRα gene promoter [19]; two years later, a second study suggested that ERRα up-regulates the endogenous SP1 gene expression by binding to its promoter region [20]. In 2006, Castet and coworkers showed by GST pull-down assay that SP1 interacts in vitro with the ERR proteins [21], although it is known that the GST protein is prone to non-specific interactions [22]. Moreover, another study showed that siRNA knockdown of SP1 reduced the transcriptional activity of ERRα at the P2 promoter of the coregulator RIP140 [23]. Taken together, these results suggest that SP1 might be a functional partner for ERRα, although the in vivo interaction remains to be further investigated. In this context, our study provides a genome-wide evidence that ERRα and SP1 are often co-occurring at the same regulatory regions and confirms the accuracy of these predictions by in situ ChIP validation.

As for ERRα, also for PGC-1α the combination of high-throughput techniques and computational analysis methods enabled us to discover a novel transcription partner: the heterodimeric protein AP-1, which is involved in a number of cellular processes including differentiation, proliferation, apoptosis and hypoxia [24, 25]. In particular, through siRNA knockdown, we could provide evidence that depletion of the AP-1 members FOS, JUN and ATF3 reduced the ability of PGC-1α to induce target gene expression. Moreover, we found that AP-1/PGC-1α targets are involved in the response to hypoxia pathway and showed that FOS and PGC-1α are co-recruited to the same regulatory sites in proximity of these target genes. Most probably, to ensure an adequate oxygen supply for oxidative metabolism in skeletal muscle cells, PGC-1α coordinates these metabolic needs through different partners: by coactivating ERRα, it induces Vegf expression and, thus, promotes vascularization; on the other hand, through AP-1, PGC-1α enhances the expression of genes involved in cell survival, oxidative stress response and reduction of inflammation (see chapter 4). In summary, we provide additional evidence of PGC-1α ability to

control and integrate different signaling pathways using a set of several transcription factor binding partners, of which AP-1 is a new component. In the future steps of this project, we could further investigate the genome-wide occupancy of AP-1 in the context of PGC-1α over-expression and verify the effects on hypoxia of this transcription factor in vitro, by keeping muscle cells under different hypoxic conditions, and in vivo, by analyzing mice performances upon treadmill running exercise in low oxygen conditions.

The identification of both SP1 and AP-1 as novel partners of ERRα and PGC-1α, respectively, was possible mainly thanks to the Principal Component Analysis (PCA) technique, that we applied on a large matrix containing all TFBSs predicted within each ChIP-Seq peak set. In the past, PCA has been extensively used for a number of biological applications and, in particular, for the analysis of genome-wide expression studies to reduce multi-dimensional data and to extract the key features explaining most of the observed variation (e.g., see [26]). In fact, as a feature extraction technique, PCA reduces the amount of information required to accurately describe a large set of data, at a relatively low computational cost. In our context, we aimed to identify the structure of binding site occurrences within ChIP-Seq peaks and, more precisely, the linear combinations of regulatory motifs that explained most of the variation in site-counts which we observed across our peaks. As described in the previous chapters, using PCA on our datasets revealed to be a very successful approach, since it allowed us to detect different types of peaks and, thus, to identify novel transcriptional partners for the chipped proteins of interest. The accuracy of our predictions was confirmed experimentally, as shown more in detail in Chapter 4 and Chapter 5.

Due to the low basal expression level of PGC-1α and since we did not have a ChIP-grade antibody for PGC-1α, we used adenoviral vectors to introduce exogenous tagged PGC-1α protein in skeletal muscle cells. In the expression array, to control for the effects of this approach, we also measured gene expression levels upon transfection with GFP adenovirus. However, despite performing the appropriate controls, this artificial condition does not completely reflect the effects of the endogenous PGC-1α protein. Moreover, using muscle cells as we did, it is not possible to study physiological processes such as angiogenesis or fiber type switch. In the light of these observations, an interesting future development of this project would be the comparison of our present results with a ChIP-Seq experiment on PGC-1α occupancy in vivo. This approach would in fact allow the investigation of the previously mentioned biological processes, as well as the study of different muscle trainings and muscle-related pathologies. Likewise, another extension of this project could be the comparison of our data from the skeletal muscle environment with PGC-1α genome-wide recruitment in other highly oxidative tissues, like

for example brain, kidney, liver and brown adipose tissue. Afterwards, an interesting experimental future step could be the study of PGC-1α post-translational modifications and of their distinct effects on the physiology of a specific tissue.

Apart from AP-1, a number of other factors were predicted as putative partners of PGC-1α, but not validated yet. A possible follow-up of this study could be the verification of these transcription factors and, thus, the determination of the global regulatory network through which PGC-1α controls skeletal muscle adaptations to internal and external stimuli. However, the transcription factors we used for our computational analysis are only a subset of all TFs present in a living cell. Indeed, a crucial factor limiting the completeness of our TFBS predictions was the fact that the sequence specificity of many TFs is still unknown. As estimated by Vaquerizas et al. in 2009 [27], the number of transcription factors in mammalian genomes is about ~1400 (~6% of the total number of protein-coding genes), hence much higher than our set of ~350 motifs. In addition to this, nuclear receptors in particular have a high degree of similarity between their recognition sequences. For example, as we did not have a motif for the ERRα paralogs ERRβ and ERRγ, it is likely that some fraction of the ERRα binding sites predicted within PGC-1α peaks might have actually been bound by either ERRβ or ERRγ. However, due to the decrease of the sequencing costs, the number of available ChIP-Seq datasets deposited in public databases is rapidly increasing. Using these ChIP-Seq peak sets, together with available SELEX data, we expect the number of known regulatory motifs to expand in the near future, allowing us to extend our motif search and to identify novel factors involved in skeletal muscle metabolism.

Moreover, to improve the accuracy of our TFBSs predictions, in the next steps of this study we could use DNase-Seq [28] to assign to each genomic region an accessibility score and reduce the number of falsely predicted peaks, as in [29]. In addition, another possibility would be to integrate our ChIP-Seq data with Hi-C deep sequencing results. This method, which is based on the Chromosome Conformation Capture technique and is able to detect chromatin interactions in the mammalian nucleus, could be used to address the limitation of long-range peak-to-gene annotation mentioned earlier and, at the same time, to get an idea of higher-order chromatin organization in skeletal muscle cells upon different treatments and in different conditions.

In conclusion, by combining high-throughput techniques with a comprehensive computational analysis, the study presented herein not only extends the repertoire of transcriptional partners for both PGC-1α and ERRα, but elucidates the mechanism through which PGC-1α is able to control the response to hypoxia in skeletal muscles. Applying our approach to other tissues or in

other physiological or pathological conditions could help to identify novel key players of relevant biological processes.

References

1. Russell, A.P., et al., *Skeletal muscle mitochondria: a major player in exercise, health and disease.* Biochimica et biophysica acta, 2014. 1840(4): p. 1276-84.
2. Shendure, J. and H. Ji, *Next-generation DNA sequencing.* Nature biotechnology, 2008. 26(10): p. 1135-45.
3. Cunningham, J.T., et al., *mTOR controls mitochondrial oxidative function through a YY1-PGC-1alpha transcriptional complex.* Nature, 2007. 450(7170): p. 736-40.
4. Murray, J. and J.M. Huss, *Estrogen-related receptor alpha regulates skeletal myocyte differentiation via modulation of the ERK MAP kinase pathway.* American journal of physiology. Cell physiology, 2011. 301(3): p. C630-45.
5. Ott, C.J., et al., *Intronic enhancers coordinate epithelial-specific looping of the active CFTR locus.* Proceedings of the National Academy of Sciences of the United States of America, 2009. 106(47): p. 19934-9.
6. Jang, W.G., et al., *Glucocorticoid receptor mediated repression of human insulin gene expression is regulated by PGC-1alpha.* Biochemical and biophysical research communications, 2007. 352(3): p. 716-21.
7. Jeong, J.H., S. Cho, and Y.K. Pak, *Sterol-independent repression of low density lipoprotein receptor promoter by peroxisome proliferator activated receptor gamma coactivator-1alpha (PGC-1alpha).* Experimental & molecular medicine, 2009. 41(6): p. 406-16.
8. Balwierz, P.J., et al., *ISMARA: automated modeling of genomic signals as a democracy of regulatory motifs.* Genome research, 2014.
9. Eisele, P.S., et al., *The peroxisome proliferator-activated receptor gamma coactivator 1alpha/beta (PGC-1) coactivators repress the transcriptional activity of NF-kappaB in skeletal muscle cells.* The Journal of biological chemistry, 2013. 288(4): p. 2246-60.
10. Brault, J.J., J.G. Jespersen, and A.L. Goldberg, *Peroxisome proliferator-activated receptor gamma coactivator 1alpha or 1beta overexpression inhibits muscle protein degradation, induction of ubiquitin ligases, and disuse atrophy.* The Journal of biological chemistry, 2010. 285(25): p. 19460-71.
11. Mootha, V.K., et al., *Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6570-5.
12. Schreiber, S.N., et al., *The estrogen-related receptor alpha (ERRalpha) functions in PPARgamma coactivator 1alpha (PGC-1alpha)-induced mitochondrial biogenesis.* Proceedings of the National Academy of Sciences of the United States of America, 2004. 101(17): p. 6472-7.
13. Liu, C., et al., *Transcriptional coactivator PGC-1alpha integrates the mammalian clock and energy metabolism.* Nature, 2007. 447(7143): p. 477-81.
14. Calvo, S., et al., *Systematic identification of human mitochondrial disease genes through integrative genomics.* Nature genetics, 2006. 38(5): p. 576-82.
15. Schreiber, S.N., et al., *The transcriptional coactivator PGC-1 regulates the expression and activity of the orphan nuclear receptor estrogen-related receptor alpha (ERRalpha).* The Journal of biological chemistry, 2003. 278(11): p. 9013-8.
16. Kallen, J., et al., *Evidence for ligand-independent transcriptional activation of the human estrogen-related receptor alpha (ERRalpha): crystal structure of ERRalpha ligand binding domain in complex with peroxisome proliferator-activated receptor coactivator-1alpha.* The Journal of biological chemistry, 2004. 279(47): p. 49330-7.
17. Xie, W., et al., *Constitutive activation of transcription and binding of coactivator by estrogen-related receptors 1 and 2.* Molecular endocrinology, 1999. 13(12): p. 2151-62.

18. Heck, S., et al., *Estrogen-related receptor alpha expression and function is associated with the transcriptional coregulator AIB1 in breast carcinoma.* Cancer research, 2009. 69(12): p. 5186-93.
19. Liu, D., et al., *Estrogen stimulates estrogen-related receptor alpha gene expression through conserved hormone response elements.* Endocrinology, 2003. 144(11): p. 4894-904.
20. Sumi, D. and L.J. Ignarro, *Sp1 transcription factor expression is regulated by estrogen-related receptor alpha1.* Biochemical and biophysical research communications, 2005. 328(1): p. 165-72.
21. Castet, A., et al., *Receptor-interacting protein 140 differentially regulates estrogen receptor-related receptor transactivation depending on target genes.* Molecular endocrinology, 2006. 20(5): p. 1035-47.
22. Walker, J.M. and R. Rapley, *Medical biomethods handbook.* 2005, Totowa, N.J.: Humana Press. xv, 644 p.
23. Nichol, D., et al., *RIP140 expression is stimulated by estrogen-related receptor alpha during adipogenesis.* The Journal of biological chemistry, 2006. 281(43): p. 32140-7.
24. Vesely, P.W., et al., *Translational regulation mechanisms of AP-1 proteins.* Mutation research, 2009. 682(1): p. 7-12.
25. Michiels, C., et al., *HIF-1 and AP-1 cooperate to increase gene expression in hypoxia: role of MAP kinases.* IUBMB life, 2001. 52(1-2): p. 49-53.
26. Wang, A. and E.A. Gehan, *Gene selection for microarray data analysis using principal component analysis.* Statistics in medicine, 2005. 24(13): p. 2069-87.
27. Vaquerizas, J.M., et al., *A census of human transcription factors: function, expression and evolution.* Nature reviews. Genetics, 2009. 10(4): p. 252-63.
28. Boyle, A.P., et al., *High-resolution mapping and characterization of open chromatin across the genome.* Cell, 2008. 132(2): p. 311-22.
29. Pique-Regi, R., et al., *Accurate inference of transcription factor binding from DNA sequence and chromatin accessibility data.* Genome research, 2011. 21(3): p. 447-55.

Acknowledgements

I would like to thank my supervisors Christoph Handschin and Erik van Nimwegen for guiding me through the adventure of an interdisciplinary PhD and for the precious help they provided me during these years. I further appreciate the critical discussions and constructive feedbacks whenever needed. I would also like to thank all current and former members of both research groups for exchanging helpful ideas and for the nice working atmosphere they shared with me. In particular, I am very grateful to Florian Geier, Anna Egger and Ivana Dinulovic for sharing opinions and for their valuable advices; to my lab mates Petra Eisele and Mario Baresic for the fruitful collaborations and for the cheerful time spent together in the lab; to Barbara Kupr for her indispensable help in completing the experiments and for the very productive collaboration. Also, I am much obliged to Marianne Liechti and Jny Wittker for their crucial help with all the

administrative issues and to Jan Welker for fixing all the IT problems. Finally, I would like to thank Tiziano, my friends and family for their unconditional support and encouragement.

Appendix

In this section, I present three manuscripts to which I contributed during the PhD.

In the first one (**Appendix 1**, *Eisele et al., J. Biol. Chem. (2013)*), I performed: the downstream steps of the microarray analysis concerning gene clustering and heat map creation; GO and KEGG enrichment analysis; binding site prediction within a region spanning 1kb around the TSS of each array gene; motif over-representation within the distinct gene clusters; motif activity response analysis (MARA).

In the second manuscript (**Appendix 2**, *Pérez-Schindler et al., Mol. Cell. Biol. (2012)*), I performed the GO analysis of NCoR1 MKO and PGC-1α mTg target gene groups and created the heat map of the genes in the array that were annotated either as “OXPHOS” or “TCA cycle”.

In the third manuscript (**Appendix 3**, *Egger et al., PLoS One (2012*)), I created the heat map of a selection of genes involved in any of the following signaling pathways: apoptosis, DNA repair, inflammation or phototransduction.

Appendix 1

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 288, NO. 4, pp. 2246–2260, January 25, 2013

The Peroxisome Proliferator-activated Receptor γ Coactivator 1α/β (PGC-1) Coactivators Repress the Transcriptional Activity of NF-κB in Skeletal Muscle Cells*[S]

Received for publication, December 4, 2012 Published, JBC Papers in Press, December 8, 2012, DOI 10.1074/jbc.M112.375253

Petra S. Eisele[‡§], **Silvia Salatino**[‡], **Jens Sobek**[¶], **Michael O. Hottiger**[§‖], **and Christoph Handschin**[‡§1]

From the [‡]*Biozentrum, Division of Pharmacology/Neurobiology, University of Basel, CH-4056 Basel, Switzerland, the* [§]*Zurich Center for Integrative Human Physiology, University of Zurich, CH-8057 Zurich, Switzerland, the* [¶]*Functional Genomics Center Zurich, ETH Zurich/University of Zurich, CH-8057 Zurich, Switzerland, and the* [‖]*Institute of Veterinary Biochemistry and Molecular Biology, University of Zurich, CH-8057 Zurich, Switzerland*

Background: Peroxisome proliferator-activated receptor γ coactivator 1 (PGC) α and PGC-1β are metabolic coactivators that are dysregulated in muscle in many chronic diseases.
Results: PGC-1α and PGC-1β differentially suppress expression of proinflammatory cytokines induced by various stimuli.
Conclusion: In muscle cells, PGC-1α and PGC-1β modulate the NF-κB pathway thus profoundly affecting inflammatory processes.
Significance: Targeting PGC-1α and PGC-1β in chronic diseases might reduce inflammation and thereby reverse disease progression.

A persistent, low-grade inflammation accompanies many chronic diseases that are promoted by physical inactivity and improved by exercise. The beneficial effects of exercise are mediated in large part by peroxisome proliferator-activated receptor γ coactivator (PGC) 1α, whereas its loss correlates with propagation of local and systemic inflammatory markers. We examined the influence of PGC-1α and the related PGC-1β on inflammatory cytokines upon stimulation of muscle cells with TNFα, Toll-like receptor agonists, and free fatty acids. PGC-1s differentially repressed expression of proinflammatory cytokines by targeting NF-κB signaling. Interestingly, PGC-1α and PGC-1β both reduced phoshorylation of the NF-κB family member p65 and thereby its transcriptional activation potential. Taken together, the data presented here show that the PGC-1 coactivators are able to constrain inflammatory events in muscle cells and provide a molecular link between metabolic and immune pathways. The PGC-1s therefore represent attractive targets to not only improve metabolic health in diseases like type 2 diabetes but also to limit the detrimental, low-grade inflammation in these patients.

A sedentary lifestyle is a strong and independent risk factor for a large number of chronic diseases including musculoskeletal, metabolic, cardiovascular and neurological disorders. These diseases have been linked to a sterile, persistent, low-grade inflammation with elevated levels of circulating cytokines like interleukin 6 (IL-6),[2] tumor necrosis factor α (TNFα), and IL-1β that often worsen disease progression (1–3). The nuclear factor κB (NF-κB) pathway is a central regulator of inflammatory processes: NF-κB activation has accordingly been associated with obesity and insulin resistance in different organs (4–6). A wide array of signals, including cytokines like TNFα and Toll-like receptor (TLR) agonists of pathogenic or dietary origin (*e.g.* excess free fatty acids (FFAs) as in obesity (7)), is able to boost NF-κB activity upon cell surface receptor binding. These ligand-receptor interactions trigger the recruitment of adaptor proteins and receptor-proximal kinases ultimately culminating in the activation of the inhibitor of NF-κB (IκB) kinase (IKK) complex. IKK subsequently phosphorylates IκB, which is then degraded by the proteasome. Decreased levels of IκB free NF-κB, thereby enabling cytosolic-nuclear translocation and ultimately transcriptional induction of a large amount of genes involved in immune function (8). The NF-κB family is comprised of 5 members RelA/p65, RelB, c-Rel, p100/p52, and p105/p50, of which the heterodimer p65/p50 is the most common form and the target of so-called "classical" NF-κB activation (8). The transcriptional activity of p65 is further modulated by post-translational modification, *i.e.* inducible phosphorylation events that affect the binding affinity to coactivators and corepressors without altering the recruitment to DNA response elements (9, 10).

Although physical inactivity clearly has a negative impact on health favoring an inflamed environment, regular, moderate

* This work was supported by grants from the Swiss National Science Foundation, Muscular Dystrophy Association USA (MDA), SwissLife "Jubiläumsstiftung für Volksgesundheit und medizinische Forschung," Swiss Society for Research on Muscle Diseases (SSEM), Swiss Diabetes Association, Roche Research Foundation, United Mitochondrial Disease Foundation (UMDF), Association Française contre les Myopathies (AFM), Gebert-Rüf Foundation "Rare Diseases" Program, and the University of Basel.
[S] This article contains supplemental Figs. S1–S6 and Tables S1 and S2.
The microarray data have been deposited in the ArrayExpress under accession number A-MEXP-1502 (array design) and E-MEXP-3676 (experimental data).
[1] To whom correspondence should be addressed: Biozentrum, University of Basel, Klingelbergstrasse 50/70, CH-4056 Basel, Switzerland. Tel.: 41-61-267-2378; Fax: 41-61-267-2208; E-mail: christoph.handschin@unibas.ch.

[2] The abbreviations used are: IL-6, interleukin 6; PGC-1α/β, peroxisome proliferator-activated receptor γ coactivator 1α/β; NF-κB, nuclear factor κB; TLR, Toll-like receptor; FFAs, free fatty acids; IκB, inhibitor of NF-κB; IKK, IκB kinase; MIP-1α (CCL3), macrophage inflammatory protein-1α; SMRT, silencing mediator of retinoic acid and thyroid hormone receptor; PPARα, peroxisome proliferator-activated receptor α.

exercise is beneficial against systemic inflammation and counteracts the development of chronic diseases (11). Besides prevention, exercise also is an effective therapeutic strategy to treat obesity, type 2 diabetes, sarcopenia, and neurodegeneration (12–14).

At the molecular level, many of the beneficial effects of exercise are mediated by the transcriptional coactivator peroxisome proliferator-activated receptor γ coactivator 1α (PGC-1α) (15). PGC-1α is transiently induced by a single bout of exercise and chronically elevated in endurance trained muscle (16). Activated PGC-1α then controls the expression of genes encoding proteins involved in mitochondrial biogenesis, oxidative phosphorylation, and other features of oxidative muscle fibers (17). Accordingly, mice with transgenic skeletal muscle-specific *Pgc-1α* overexpression perform better in endurance tests and display a switch toward oxidative type I and type IIA fibers (18, 19). The increased fitness of healthy transgenic animals translates into improvement of symptoms when *Pgc-1α* is expressed in the context of different muscle wasting conditions as shown for Duchenne muscular dystrophy, sarcopenia, a mitochondrial myopathy and denervation- or lovastatin-induced fiber atrophy (20–22).

Inversely, a skeletal muscle-specific deletion of the *Pgc-1α* gene facilitates a type IIB and type IIX fiber type switch, reduces exercise performance, and promotes muscle fiber damage (23). Furthermore, loss of *Pgc-1α* results in elevated levels of proinflammatory factors locally in muscle as well as systemically (24). These findings suggest an anti-inflammatory role for PGC-1α: in fact, in skeletal muscle of diabetic patients, PGC-1α levels negatively correlate with *Il-6* or *Tnfα* levels independent of body mass index (24).

Pgc-1β, a closely related member of the *Pgc-1* gene family, also exhibits dysregulated expression in skeletal muscle of diabetic patients and thereby contributes to the mitochondrial dysfunction observed in type 2 diabetes (25). Although both PGC-1s share the ability to boost oxidative metabolism, PGC-1β is not regulated by exercise and primarily drives the formation of type IIX fibers (26). Interestingly, PGC-1β is required for alternative activation of and mitochondrial reactive oxygen species production in macrophages (27, 28); an immunomodulatory role in skeletal muscle has, however, not been described so far.

Based on these observations, we now tested the hypothesis that the PGC-1 coactivators exert anti-inflammatory effects in muscle. More precisely, we explored if PGC-1α and PGC-1β are able to modify cytokine expression upon exposure of muscle cells to different inflammatory stimuli like TNFα, TLR agonists, and FFAs. We found that the PGC-1s repress the transcriptional activity of p65 and thereby modulate the NF-κB signaling pathway. These data represent a prime example of cross-talk between metabolic and immune pathways with important implications for skeletal muscle function.

EXPERIMENTAL PROCEDURES

Cell Culture and Treatments—The mouse skeletal muscle cell line C2C12 was maintained below confluence in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal calf serum and 1× penicillin/streptomycin (Invitrogen). For differentiation into myotubes, growth medium was exchanged for DMEM supplemented with 2% horse serum (Invitrogen) for at least 3 days. *Pgc-1α*, *Pgc-1β*, and *Gfp* were overexpressed from recombinant adenoviral constructs 48 h prior to treatment. Stimulation with TNFα (Sigma) and TLR agonists (Invivogen) in growth or differentiation medium lasted for 2 h unless otherwise stated. Concentrations were as follows: TNFα, 10 ng/ml; PAM3CSK4, 1 μg/ml (TLR1/2 agonist); HKLM, 10^8 cells/ml (TLR2 agonist); poly(I:C), 25 μg/ml (TLR3 agonist); *Escherichia coli* K-12 LPS, 1 μg/ml (TLR4 agonist); *Salmonella typhimurium* flagellin, 1 μg/ml (TLR5 agonist); FSL1, 1 μg/ml (TLR6/2 agonist); ssRNA40, 1 μg/ml (TLR8 agonist); and ODN18266, 5 μM (TLR9 agonist). FFA (Sigma) were dissolved in ethanol and further diluted to 1 mM final concentration in DMEM containing 2% fatty acid- and endotoxin-free bovine serum albumin (Sigma); FFA treatment lasted for 16 h in serum-free medium. The protein phosphatase inhibitor okadaic acid (Sigma, 250 nM) was present 30 min prior to and during treatment with TNFα where indicated, whereas control samples were incubated with vehicle (DMSO, 0.04%) alone for equal times. The PPARα inhibitor MK 886 (Tocris Bioscience, 20 μM) was present 24 h prior to and during treatment with TNFα, where indicated, whereas control samples were incubated with vehicle (DMSO, 0.02%) alone for equal times.

Semiquantitative Real-time PCR—RNA was isolated from treated C2C12 cells using TRIzol (Invitrogen) and residual DNA contamination was removed by DNase I digestion (Invitrogen). 1 μg of RNA was reverse transcribed with SuperScript II (Invitrogen) and the resulting cDNA was used as template for RT-PCR. To detect relative expression levels, cDNA was amplified with the SYBR Green Master mix (Applied Biosystems) and analyzed on a StepOnePlus RT-PCR System (Applied Biosystems). The respective primer pairs are listed in supplemental Table S1. All values are normalized to the expression of TATA-box binding protein and expressed as fold-induction over the untreated control condition.

ELISA—To determine cytokine concentrations in cell culture supernatants, sandwich immunoassays against IL-6 were performed according to the manufacturer's instructions (Quantikine, R&D Systems).

NF-κB Customized Array Analysis—Differentiated C2C12 cells overexpressing *Pgc-1α*, *Pgc-1β*, or *Gfp* were treated with TNFα for 2 h and RNA was extracted (NucleoSpin RNA II, Macherey-Nagel). The custom array analysis was subsequently performed as previously described (29). Briefly, RNA was processed with the Amino Allyl MessageAMP II aRNA Amplification Kit (Ambion) according to the manufacturer's instructions to yield aminoallyl-modified aRNA that was further coupled to Cy3 or Cy5 dyes, respectively. Biological duplicates were labeled cross-wise to control for dye effects. Pairs of samples were hybridized on a total of 14 slides containing probes for 524 genes included in the array as described elsewhere (29, ArrayExpress A-MEXP-1502). Significantly regulated genes as determined by analysis of variance ($p \leq 0.01$) were subject to cluster analysis. The 3 resulting clusters were the source of further KEGG (Kyoto Encyclopedia of Genes and Genomes) and GO (Gene Ontology) terms enrichment analysis with FatiGO (30) and considering the adjusted p value ≤ 0.05 as significant. To

predict the relevant transcription factors for regulation of each cluster, a motif search within a region spanning +200 bp to −800 bp around the transcription start site of each gene in the array was performed using MotEvo (31) in combination with a nonredundant set of position weight matrices mainly derived from JASPAR and TRANSFAC. Overrepresented motifs were retrieved by comparing binding site occurrence in each cluster to the occurrence in the whole array and considered significant with a Z score ≥2. Microarray data were further analyzed with MARA (Motif Activity Response Analysis (32)) to predict the most important transcription factors that contribute to the changes in gene expression in our set of samples (independent of clustering).

Dual-luciferase Reporter Gene Assays—90–95% confluent C2C12 myoblasts were transfected using Lipofectamine 2000 (Invitrogen) with plasmids containing firefly luciferase under the control of 3 NF-κB sites (wild type, WT) or a construct with mutated sites as control in combination with *Renilla* luciferase as internal control. Cells were then either treated with TNFα/TLR agonists for 2 h, or co-transfected with p65, increasing amounts of *Pgc-1α/β*, and empty vector (pcDNA3.1) to keep total DNA concentration constant. Cells were assayed 24 h post-transfection with the Dual Luciferase Reporter Assay System (Promega) according to the manufacturer's protocol. Briefly, cells were lysed and firefly and *Renilla* luciferase activities were determined on a Centro LB 960 luminometer (Berthold Technologies). Firefly-derived luciferase values were normalized to *Renilla* activity and results expressed as ratio of WT to mutated reporter with the control condition being arbitrarily set to 1.

TransAM NF-κB DNA-binding Assays—Nuclear extracts from myotubes were prepared by swelling the cells in hypotonic buffer (20 mM Hepes, pH 7.5, 5 mM NaF, 10 μM Na_2MoO_4, 0.1 mM EDTA) and breaking the cytoplasmic membrane with Nonidet P-40 substitute (0.5%). After centrifugation, the residual nuclear pellet was resuspended in Complete Lysis Buffer to dismantle nuclear membrane integrity and obtain the nuclear extract as supernatant of the final, fast centrifugation step. To determine the amount of active NF-κB in these nuclear extracts, TransAM assays (Active Motif) were performed following the manufacturer's instructions. In brief, 10 μg of nuclear extracts were added to wells coated with oligonucleotides containing the NF-κB consensus sequence (5′-GGGACTTTCC-3′). These oligonucleotides trap active transcription factors, which then are detected in an ELISA-like assay with specific primary antibodies for the different NF-κB subunits and a secondary horseradish peroxidase (HRP)-conjugated antibody. A colorimetric reaction accordingly determines the amount of NFκB-DNA binding. Results are expressed in arbitrary units minus the respective blank value, and normalized to Ctrl conditions.

Western Blotting—After washing away residual culture medium with PBS, cells were scraped off the dish in lysis buffer (20 mM Tris-HCl, pH 7.5, 150 mM NaCl, 1 mM EDTA, 1 mM EGTA, 1% Triton X-100, 2.5 mM $Na_4P_2O_7$, 1 mM β-glycerophosphate, 1 mM Na_3VO_4) supplemented with protease (complete tablets, Roche) and phosphatase inhibitors (phosphatase inhibitor mixture 2, Sigma) and kept on ice for 10 min with occasional vortexing. Centrifugation at 10,000 × g yielded the protein lysate, which was subjected to SDS-PAGE. The resolved proteins were transferred to a nitrocellulose membrane (Whatman) on a semi-dry electroblotting system (Thermo Scientific) with 1.5 mA/cm^2 and transfer was confirmed by Ponceau S staining (0.1% in 5% acetic acid). Proteins were detected with the following primary antibodies: anti-p65, anti-Pp65(Ser-536) (both Cell Signaling Technology), anti-p105/p50 (Abcam), anti-p100/p52, anti-RelB, anti-cRel, anti-IκBα, anti-IκBβ, anti-IKKα, anti-IKKγ, anti-α-tubulin as loading control (all Cell Signaling Technology) and an anti-rabbit IgG HRP-coupled secondary antibody (Dako). Bands were finally visualized by autoradiography with ECL substrate (Pierce) and density was quantified with ImageJ software.

Statistical Analysis—Data were analyzed with Student's t test using $p \leq 0.05$ as the significance threshold.

RESULTS

PGC-1α and PGC-β Differentially Suppress Proinflammatory Cytokine Expression and Secretion—To delineate the effect of PGC-1s on inflammatory processes in muscle, myotubes overexpressing either adenovirally encoded *Pgc-1α* or *Pgc-1β* were compared with *Gfp*-expressing control cells treated with different inflammatory stimuli. *Pgc-1* coactivator overexpression was high and functional as evident from the induction of the known target gene medium-chain acyl-coenzyme A dehydrogenase (supplemental Fig. S1, *A* and *B*). Ectopic, recombinant TNFα strongly induced the gene expression of proinflammatory markers *Il-6*, *Tnfα*, and macrophage inflammatory protein-1α (*Mip-1α/Ccl3*). Although PGC-1α did not negatively affect basal levels of these proinflammatory cytokines, it diminished their induction by TNFα yielding lower levels than in control cells (Fig. 1*A*). PGC-1β suppressed both basal and TNFα-induced expression of IL-6, TNFα, and MIP-1α (Fig. 1*B*). As a consequence of the PGC-1-mediated repression in gene expression, lower levels of secreted IL-6 protein were observed in culture media after TNFα treatment (Fig. 1, *C* and *D*).

Next, we examined the effect of PGC-1s on TLR stimulation in muscle cells. First, we assessed the expression pattern of the *Tlr*s in our experimental system and found that *Tlr*1, *Tlr*2, *Tlr*3, *Tlr*4, *Tlr*6, and at lower levels *Tlr*5 were detectable in the muscle cells (Fig. 2*A*). To study the activity of the TLRs, selective agonists for TLR2, TLR3, TLR4, TLR5, TLR8, TLR9, or the TLR1/2 and TLR6/2 heterodimers were applied. Of those, only TLR1/2, TLR4, and TLR6/2 activators consistently elevated expression of *Il-6* and *Tnfα* (Fig. 2*B*). Moreover, the same compounds increased luciferase reporter gene activity controlled by 3 repeats of a minimal NF-κB DNA-response element (Fig. 2*C*). For the subsequent experiments, agonists for TLR1/2, TLR4, and TLR6/2, the active TLRs in muscle cells, were utilized and compared with the TLR3 activator that was used as a negative, internal control. TLR1/2, TLR4, and TLR6/2 agonists all induced expression of *Il-6* and *Tnfα* in control cells as expected (Fig. 2*D*). Interestingly, PGC-1α effectively repressed TLR-mediated TNFα expression by all of the active agonists, whereas *Il-6* expression was only reduced by PGC-1α in the TLR1/2 and TLR4 agonist treatment (Fig. 2*D*). PGC-1α, furthermore, low-

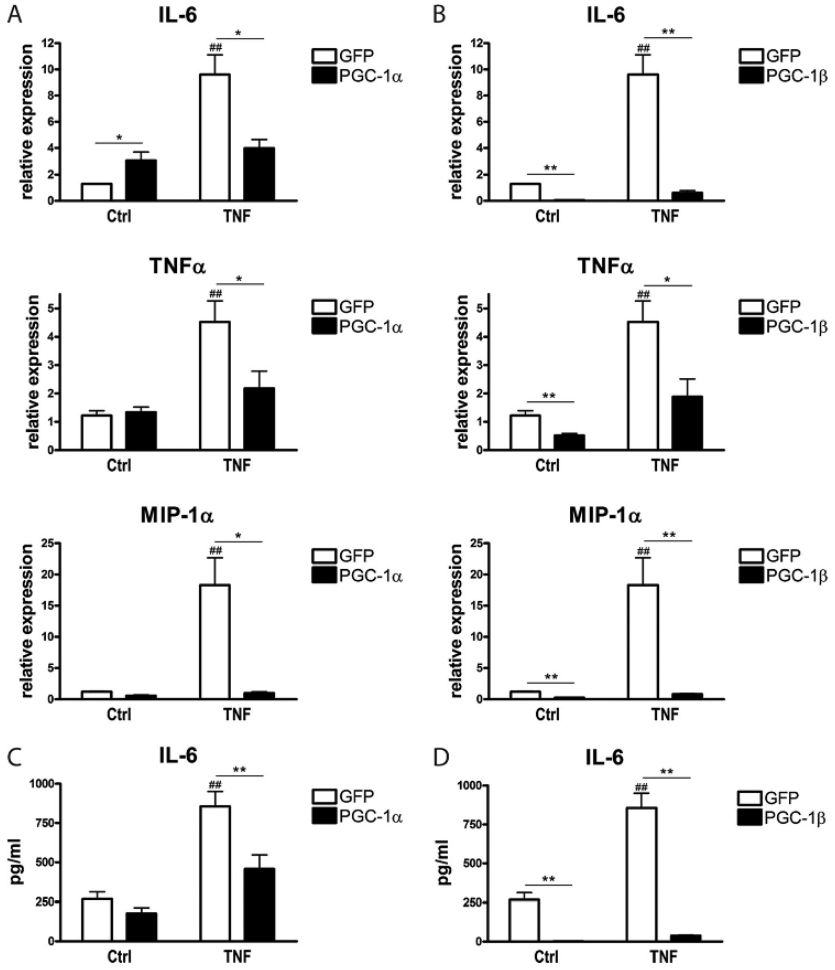


FIGURE 1. **PGC-1α and PGC-1β suppress TNFα-induced proinflammatory cytokines.** *A-D*, differentiated C2C12 myotubes overexpressing *Pgc-1α* and *Gfp* (*panels A* and *C*) or *Pgc-1β* and *Gfp* (*panels B* and *D*) were treated with TNFα for 2 h. Expression of proinflammatory cytokines was determined by real-time PCR (*panels A* and *B*) and release of IL-6 into the medium was quantified by ELISA (*panels C* and *D*). Values represent the at least 3 independent experiments, mean ± S.E. ##, $p \leq 0.01$ GFP TNF *versus* GFP control (*Ctrl*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

ered basal IL-6 protein secretion, but not the elevated IL-6 secretion caused by TLR agonist stimulation (Fig. 2*E*). In contrast, PGC-1β significantly decreased *Il-6* and *Tnfα* gene expression both in the basal state and after activation of TLR1/2, whereas no effect was observed in the TLR4 or TLR6/2 agonist-treated cells (Fig. 2*F*). Nevertheless, however, IL-6 secretion into culture medium was strongly repressed by PGC-1β in all of the different experimental conditions (Fig. 2*G*).

As a third inflammatory stimulus besides TNFα and TLR agonists, the FFA palmitic acid ($C_{16}H_{32}O_2$), oleic acid ($C_{18}H_{34}O_2$), myristic acid ($C_{14}H_{28}O_2$), stearic acid ($C_{18}H_{36}O_2$), linoleic acid ($C_{18}H_{32}O_2$), and elaidic acid ($C_{18}H_{34}O_2$) were administered to muscle cells to mimic the lipid overload that is associated with disease progression in the metabolic syndrome. As expected, saturated fatty acids, in particular stearic and palmitic acid and to a lesser extent the shorter chain species, produced a strong proinflammatory response as indicated by the transcriptional induction of the *Il-6*, *Tnfα*, and *Mip-1α* genes (Fig. 3*A*). In contrast, mono- or polyunsaturated fatty acids did not alter proinflammatory cytokine expression (Fig. 3*A*); in fact, oleic acid was even able to reverse the negative effects of palmitic acid (Fig. 3*A*) as previously described (33). Strikingly, PGC-1α potently inhibited the increase in *Il-6*, *Tnfα*, and *Mip-1α* gene expression mediated by palmitic, myristic, and stearic acid (Fig. 3*A*). Likewise, PGC-1β also efficiently blocked the FFA-induced elevation of *Il-6* and *Mip-1α* transcript levels (Fig. 3*B*). Interestingly, however, transcriptional elevation of *Tnfα* by palmitic, myristic, and stearic acid was unaffected by overexpression of *Pgc-1β* (Fig. 3*B*).

PGC-1α and PGC-1β Target the NF-κB Pathway to Suppress Inflammation—To examine the mechanisms behind the repressive effect of the PGC-1 coactivators on TNFα-induced proinflammatory cytokines, a customized array designed to represent the most important inflammatory genes and other NF-κB targets (29) was employed. Of the 524 genes that are

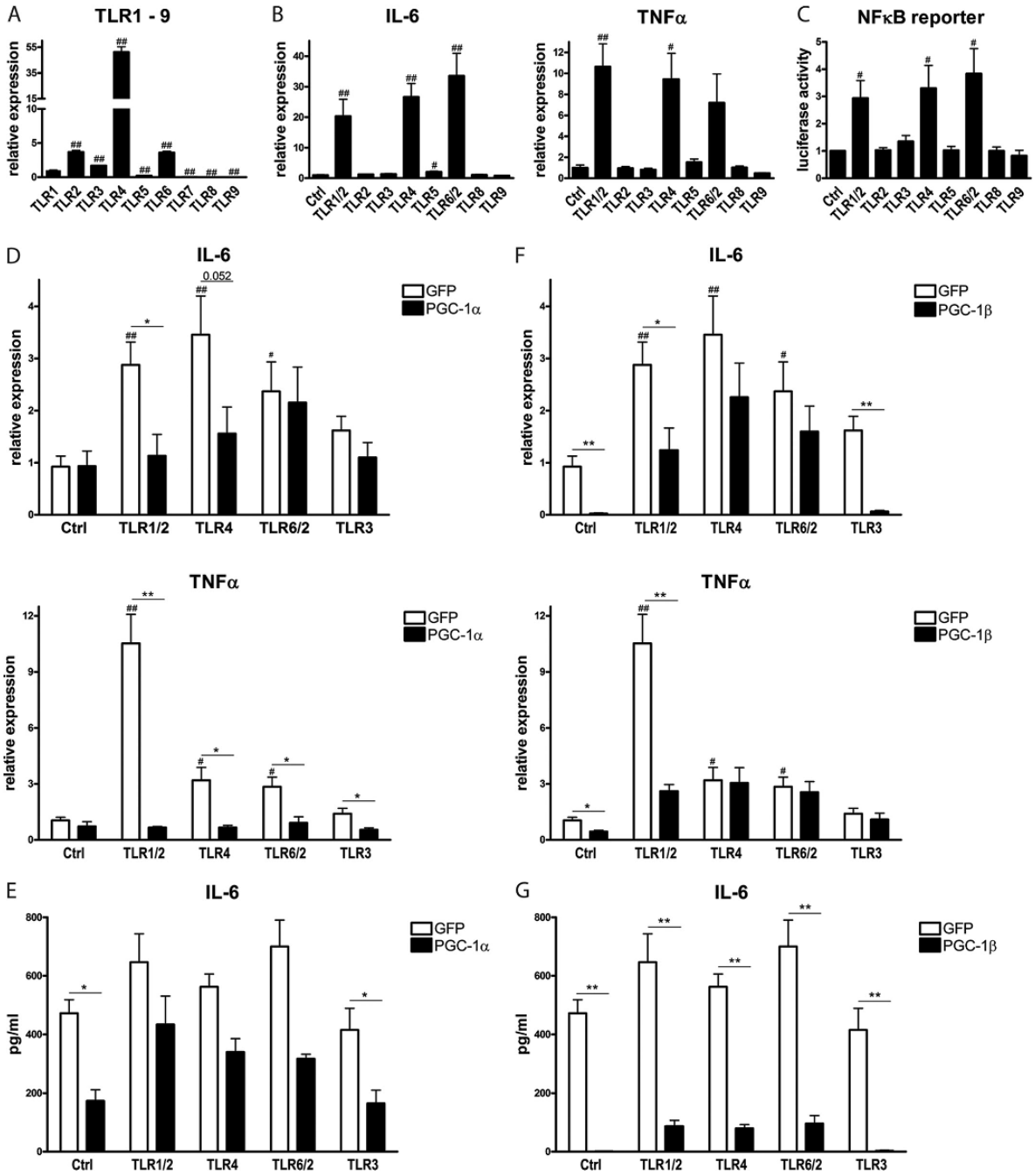


FIGURE 2. **PGC-1α and PGC-1β differentially suppress TLR agonist-induced proinflammatory cytokines.** *A*, differentiated C2C12 myotubes were tested for mRNA expression of *Tlr1–9*. Expression values are arbitrarily normalized to those of *Tlr1*. *B*, C2C12 cells were differentiated and treated with different TLR agonists for 2 h. Expression of *Il-6* and *Tnfα* was assessed by real-time PCR. *C*, C2C12 cells were transfected with an NF-κB reporter construct (or a mutated reporter construct as control) and treated with different TLR agonists for 2 h. Luciferase activity was determined and is expressed as ratio of WT to mutated reporter gene values. *D–G*, differentiated C2C12 myotubes overexpressing *Pgc-1α* and *Gfp* (*panels D* and *E*) or *Pgc-1β* and *Gfp* (*panels F* and *G*) were treated with agonists for TLR1/2, TLR4, TLR6/2, and TLR3 for 2 h. Expression of proinflammatory cytokines was determined by real-time PCR (*panels D* and *F*) and release of IL-6 into the medium was quantified by ELISA (*panels E* and *G*). Values represent at least 3 independent experiments, mean ± S.E. #, $p \leq 0.05$; ##, $p \leq 0.01$ TLR agonist treated *versus* control (*Ctrl*) (*A*, *B*, and *D-G*) or TLR expression *versus* TLR1 (*C*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

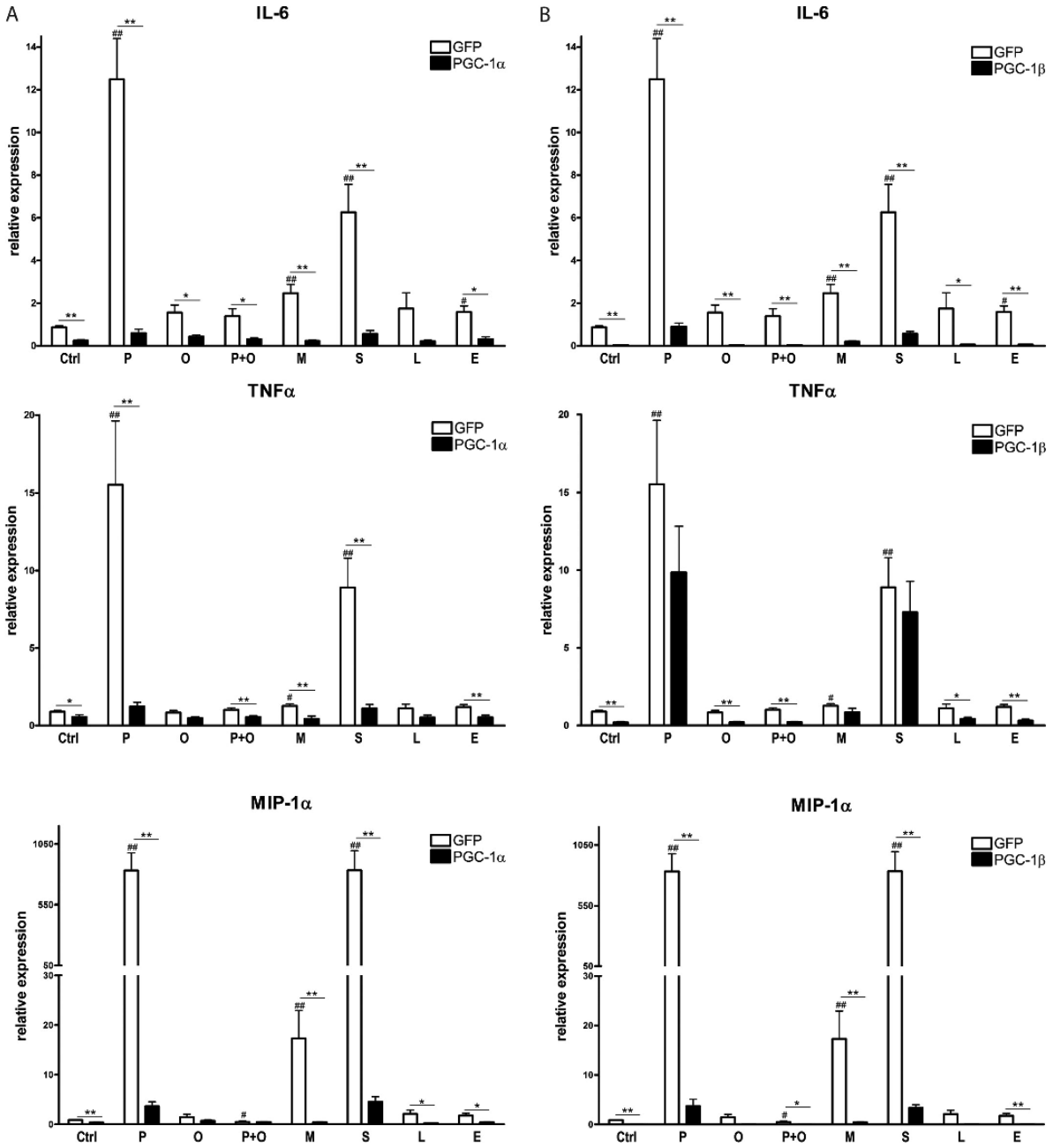


FIGURE 3. **PGC-1α and PGC-1β differentially suppress fatty acid-induced proinflammatory cytokines.** *A* and *B*, differentiated C2C12 myotubes overexpressing *Pgc-1α* and *Gfp* (*panel A*), or *Pgc-1β* and *Gfp* (*panel B*) were treated with different fatty acids (*P* = palmitic acid, *O* = oleic acid, *M* = myristic acid, *S* = stearic acid, *L* = linoleic acid, *E* = elaidic acid) for 16 h. Expression of proinflammatory cytokines was determined by real-time PCR. Values represent at least 3 independent experiments, mean ± S.E. #, $p \leq 0.05$; ##, $p \leq 0.01$ GFP FFA *versus* GFP control (*Ctrl*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

present on the array, 55 genes were found to be differentially regulated by TNFα treatment and/or *Pgc-1* overexpression ($p < 0.01$) and thus further analyzed. These 55 genes were grouped into 3 clusters (Fig. 4*A*): cluster 1 contained 21 genes that were up-regulated by PGC-1α and PGC-1β including vascular endothelial growth factor (*Vegf*), a known PGC-1α target (34). The 13 and 21 genes in clusters 2 and 3, respectively, were repressed by PGC-1α and PGC-1β. Importantly, genes in cluster 3 were TNFα-inducible, whereas expression of the genes in cluster 2 was not modulated by TNFα treatment (Fig. 4*A*). To validate the results of the microarray, three representative genes from each cluster were chosen and their expression ana-

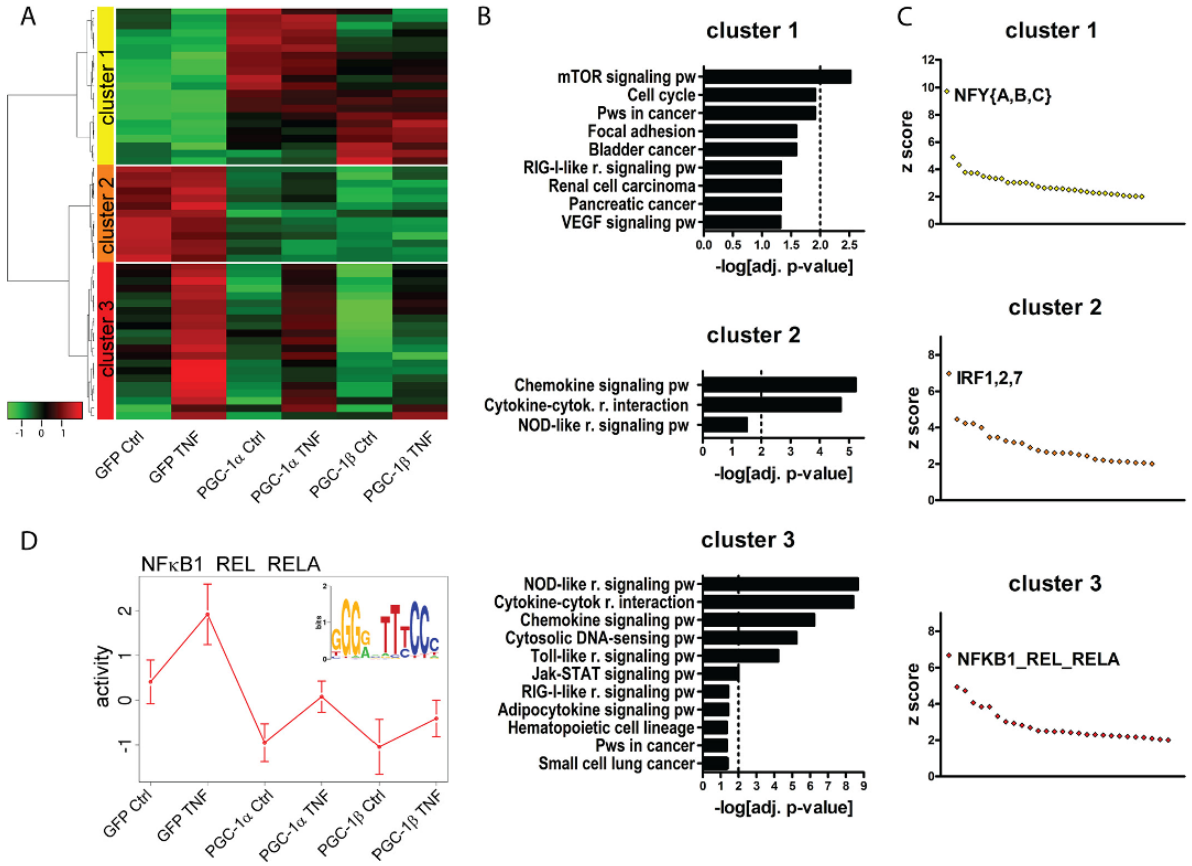


FIGURE 4. **Inflammatory pathways and NF-κB activity are suppressed by PGC-1s.** Differentiated C2C12 myotubes overexpressing *Pgc-1α*, *Pgc-1β*, and *Gfp* were treated with TNFα for 2 h and subsequently subjected to a customized microarray analysis. *A*, significantly regulated genes ($p \leq 0.01$) were clustered and are depicted as a heat map. *B*, KEGG pathways enriched within each cluster (adjusted *p* value ≤ 0.05) are shown, the *dotted line* indicates an adjusted *p* value of 0.01. Abbreviations: *cytok.* = cytokine, *pw* = pathway, *r.* = receptor). *C*, motifs overrepresented in the promoters of each cluster were identified and their *Z* score distribution plotted. For a complete list see supplemental Table S2A. *D*, activity plot of NF-κB (top scoring transcription factor motif) over different conditions as predicted by MARA (motif activity response analysis) and the corresponding sequence logo of the position weight matrix.

lyzed by real-time PCR. Mitochondrial translational initiation factor 2 (*Mtif2*), *Vegfα*, and protein arginine methyltransferase 1 (*Prmt1*) from cluster 1 were indeed induced (supplemental Fig. S2*A*), whereas chemokine (C-X-C motif) ligand 12 (*Cxcl12*), complement component 2 (*C2*), and E2F transcription factor 2 (*E2f2*) from cluster 2 (supplemental Fig. S2*B*) and chemokine (C-C motif) ligand 2 (*Cd2*), *Cd7*, and *Cxcl1* from cluster 3 (supplemental Fig. S2*C*) were repressed by PGC-1 coactivators confirming the results of the microarray. Furthermore, TNFα inducibility of genes in cluster 3 was verified.

Functionally, the genes in cluster 1 were enriched in only one Gene Ontology (GO) term compared with 18 significant terms in cluster 2 and 96 terms in cluster 3 (of which only 46 terms with a *p* value ≤ 0.01 are shown) (supplemental Fig. S3). 12 of the GO categories in cluster 2 and 26 categories in cluster 3 are related to inflammation and immunity. Importantly, the inflammation-related GO terms comprise the top 10 and 15 ranking categories in clusters 2 and 3, respectively (supplemental Fig. S3). Likewise, all 3 KEGG pathways assigned to cluster 2 and the top 9 KEGG pathways out of 11 of cluster 3 are related to inflammatory signaling (Fig. 4*B*). These results suggest that the PGC-1s are able to repress inflammatory processes in muscle cells.

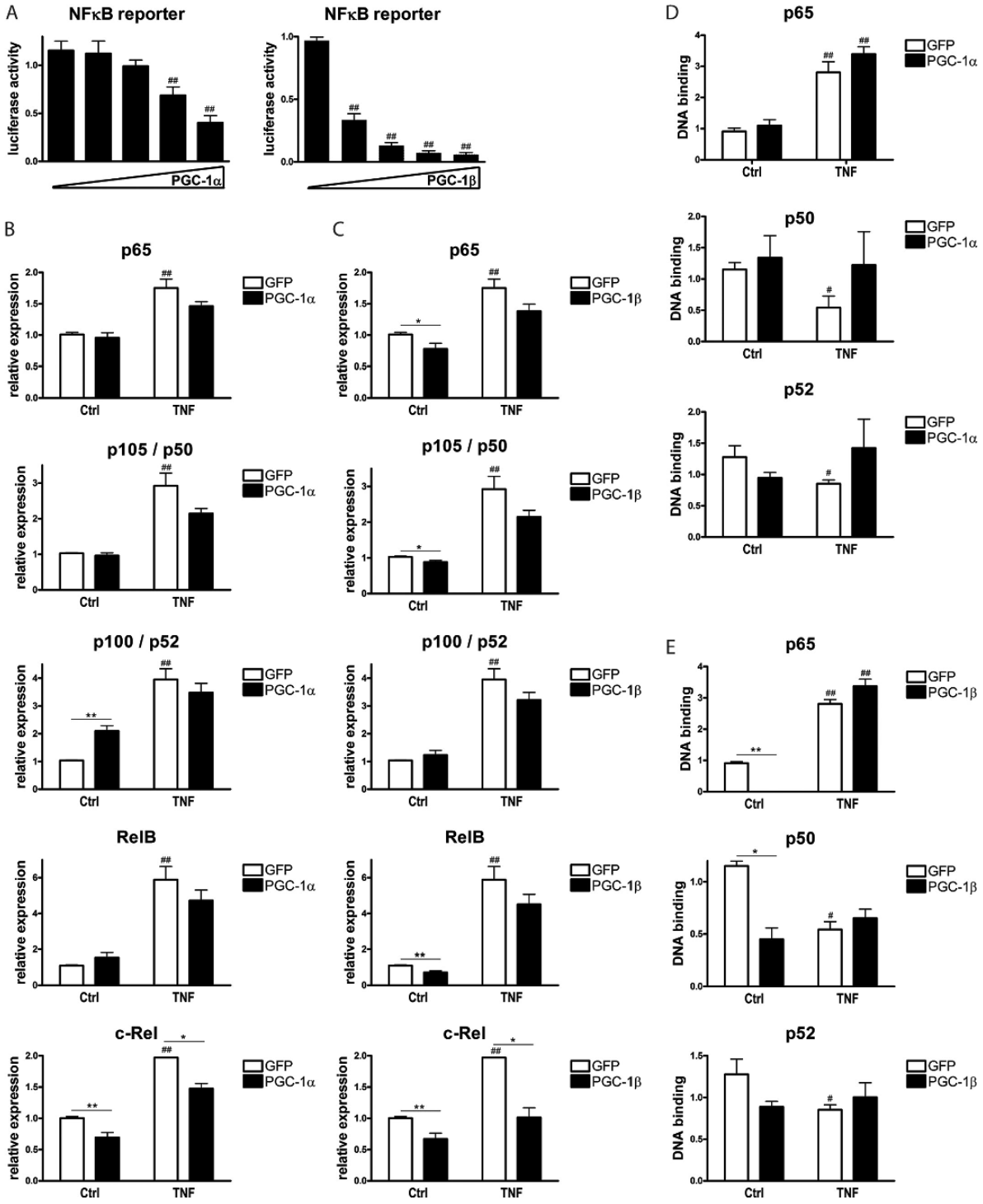
To predict which transcription factors are involved in the regulation of each cluster, promoter regions were analyzed in regard to their motif composition. Motifs overrepresented in each of the clusters are listed in supplemental Table S2A. As background for this analysis, all promoters of the microarray were used, thereby eliminating any putative bias that might have been introduced with the specific choice of genes in the customized microarray. The majority of predicted binding sites were unique to one cluster, *i.e.* 26 of 33 motifs in cluster 1, 20 of 26 motifs in cluster 2, and 21 of 28 motifs in cluster 3 (supplemental Table S2*B*). Intriguingly, one distinct transcription factor binding motif stood out as the clear top ranking candidate in each of the three clusters based on *Z* score, namely NFY{A,B,C} in cluster 1 (*Z* score of 9.74 compared with the second ranking Z score of 4.9 for GFI1), IRF1, -2, and -7 in cluster 2 (*Z* score of 6.98 compared with 4.46 for FOX{F1,F2,J1}), and NFκB1_REL_RELA in cluster 3 (Z score of 6.67 compared with

4.93 for AR) (Fig. 4*C* and supplemental Table S2*A*). These predictions indicate that the NF-κB pathway is a likely target for the PGC-1s to suppress TNFα-inducible inflammatory gene expression represented in cluster 3. Accordingly, NF-κB was also the highest scoring motif in the motif activity response analysis of the whole array and thus the most likely transcription factor to modulate TNFα- and PGC-1-dependent gene expression (Fig. 4*D*).

PGC-1β Reduces p65 and p50 Expression Levels—To experimentally validate the biocomputational prediction of PGC-1-mediated repression of NF-κB signaling as the central mechanism for the anti-inflammatory effect of the PGC-1 coactivators on TNFα-inducible genes, reporter gene assays were performed with a construct containing the luciferase gene under control of 3 NF-κB DNA response elements. The activity of the reporter gene construct in response to TNFα treatment resembles endogenous *Il-6* and *Tnfα* gene expression (supplemental Fig. S1, *C* and *D*). Increasing amounts of co-transfected *Pgc-1α* or *Pgc-1β* progressively inhibited transcription from a NFκB-responsive promoter that is activated by exogenous p65 (Fig. 5*A*).

To elucidate the mechanism by which the PGC-1 coactivators repress NF-κB activity even on minimal NF-κB response element-driven gene expression, we first examined the expression levels of the different NF-κB family members before and after TNFα treatment and *Pgc-1* overexpression, respectively (Fig. 5, *B* and *C*). TNFα led to a significant increase in transcript levels of all 5 NF-κB isoforms (*p65*, *p105*/*p50*, *p100*/*p52*, *RelB*, and *c-Rel*) in *Gfp*-infected control cells. PGC-1α did not change the levels of the canonical isoforms *p65* and *p105*/*p50* as well as *RelB* in the basal, vehicle-treated cells, whereas an induction of *p100*/*p52* was observed suggesting a switch toward the noncanonical NF-κB pathway. In contrast to the first 4 NF-κB isoforms, *c-Rel* expression was clearly reduced by *Pgc-1α* overexpression (Fig. 5*B*). After TNFα treatment, the levels of all NF-κB isoforms tended to be lower in PGC-1α overexpressing cells, however, only reaching statistical significance in the case of *c-Rel* gene expression (Fig. 5*B*). In striking contrast to the PGC-1α-mediated effect, PGC-1β suppressed *p65* and *p105*/*p50* as well as *RelB* and *c-Rel* gene expression in the basal state (Fig. 5*C*). The transcript levels of all 5 NF-κB isoforms tended to be lower in *Pgc-1β* overexpressing cells after TNFα treatment; similar to *Pgc-1α* overexpressing cells, this repression was, however, only significant for *c-Rel* gene expression (Fig. 5*C*). Therefore, whereas the basal repression of *p65* and *p105*/*p50* could contribute to the strong repression of inflammatory gene expression by PGC-1β in nonstimulated cells, it is unlikely that the small differences in expression levels of the different NF-κB isoforms after TNFα treatment underlie the profound effects of PGC-1s on proinflammatory cytokine expression in stimulated muscle cells. Thus, to further elucidate the molecular mechanism underlying this observation, we next assessed the DNA binding capability of all NF-κB family members in the context of TNFα treatment and *Pgc-1* overexpression (Fig. 5, *D* and *E*). Of the 5 NF-κB isoforms, RelB and c-Rel were undetectable in nuclear extracts from C2C12 myotubes in the TransAM assay and are thus unlikely to play a major role in the regulation of proinflammatory cytokine expression in our experimental context of muscle cells as hypothesized in other publications (35). The DNA binding of p50 and p52 was detectable although low (Fig. 5, *D* and *E*). In contrast, recruitment of p65 to DNA response elements was substantial in nonstimulated cells and further elevated after TNFα treatment. Interestingly however, DNA binding of p65 with and without TNFα treatment, respectively, was not changed by PGC-1α (Fig. 5*D*), whereas PGC-1β strongly inhibited binding of p65 and p50 to DNA in the basal, vehicle-treated muscle cells (Fig. 5*E*). Comparable with *Pgc-1α* overexpression, DNA binding of any of the NF-κB isoforms was not affected by ectopic PGC-1β in TNFα-stimulated cells (Fig. 5*E*). It is thus conceivable that both low *p65*/*p50* expression as well as reduced DNA binding account for the PGC-1β-mediated reduction in proinflammatory cytokine expression compared with vehicle-treated control cells. In contrast, however, the diminished levels of these cytokines upon TNFα treatment in both *Pgc-1α*- and *Pgc-1β*-overexpressing cells can neither be attributed to changes in NF-κB expression nor to modulation of the NF-κB protein binding capability to DNA response elements.

PGC-1α and PGC-1β Diminish the Transcriptional Activity of p65—Based on our data implying alternative molecular mechanisms distinct from transcriptional regulation or DNA binding of NF-κB to underlie the repressive action of the PGC-1 coactivators on the activity of this transcription factor, we next studied upstream signaling and the post-translational modification of p65 that influence the transcriptional activity of NF-κB. First, we examined the possibility that high levels of IκBs after TNFα treatment could account for lowered cytokine expression. However, the relative amount of IκBβ protein was not different between conditions (Fig. 6*A*). *IκBα*, an NF-κB target gene, was accordingly increased on the protein level by TNFα treatment and reduced in muscle cells overexpressing *Pgc-1α* or *Pgc-1β* (Fig. 6*A* and supplemental Fig. S4, *A* and *B*), similar to other NF-κB targets such as *Il-6* and *Tnfα* (Fig. 1) excluding the possibility of IκB-mediated repression. Subsequently, IKK protein levels were assessed: IKKβ was not and IKKγ was barely detectable (Fig. 6*A*). In contrast, IKKα was slightly increased by TNFα treatment; this effect was abrogated in cells overexpressing *Pgc-1α* or *Pgc-1β* (Fig. 6*A*). IKKα is one of the protein kinases that is able to phosphorylate p65 at serine 536. Importantly, the phosphorylation status of p65 at serine 536 affects the transcriptional activity of NF-κB even when bound to DNA response elements (10). In Western blot analyses of total and phosphorylated p65 protein, a small but significant increase in total p65 protein levels was observed after TNFα treatment in control cells (Fig. 6, *A* and *B*). *Pgc-1α* overexpression did not affect basal levels of p65, whereas *Pgc-1β* diminished total p65 protein expression in this context, as expected based on the reduced mRNA expression of *p65* in *Pgc-1β* overexpressing muscle cells (Fig. 5*C*). Strikingly, however, both PGC-1α and PGC-1β reduced TNFα-mediated phosphorylation of p65 by about 50% (Fig. 6, *A* and *B*). None of the other NF-κB family members underwent regulation by TNFα on the protein level. Interestingly, *Pgc-1α* overexpression resulted in elevated protein levels of p105, p100, p52, and


A
NFκB reporter
luciferase activity
PGC-1α
NFκB reporter
luciferase activity
PGC-1β
B
p65
relative expression
GFP
PGC-1α
Ctrl
TNF
p105 / p50
p100 / p52
RelB
c-Rel
C
p65
GFP
PGC-1β
p105 / p50
p100 / p52
RelB
c-Rel
D
p65
DNA binding
p50
p52
E
p65
p50
p52

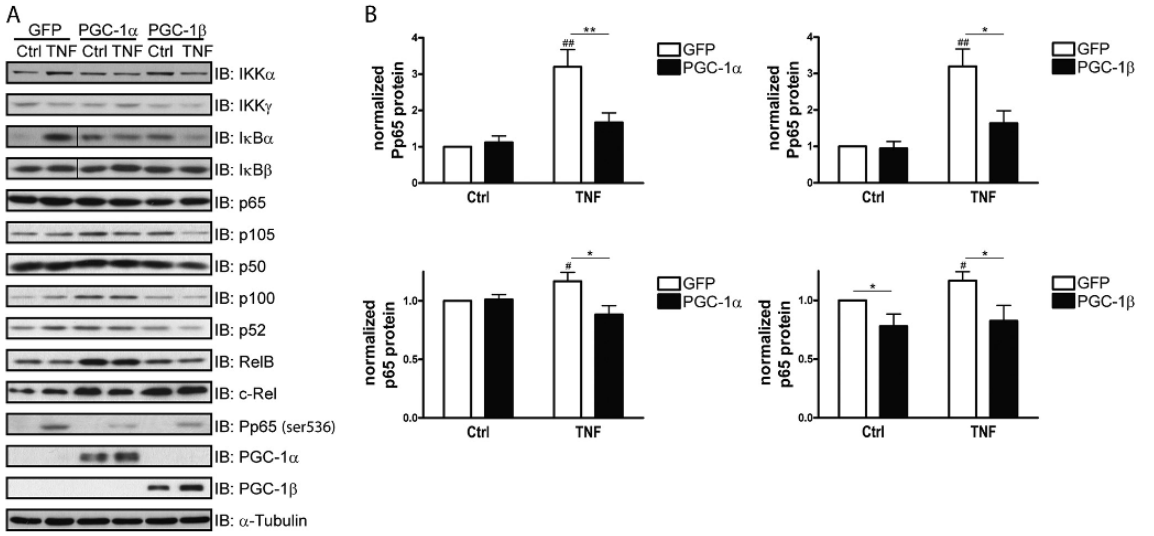


FIGURE 6. **PGC-1s diminish p65 phosphorylation at serine 536.** Differentiated C2C12 myotubes overexpressing *Pgc-1α*, *Pgc-1β*, and *Gfp* were treated with TNFα for 2 h. *A*, protein abundance of NF-κB family members, pathway components, phospho-p65 (Ser-536), PGC-1α, PGC-1β, and α-Tubulin was assessed by immunoblotting (*IB*). *B*, quantification of phospho-p65 (Ser-536) and total p65 protein levels, normalized to α-tubulin. Values represent at least 3 independent experiments, mean ± S.E. #, $p \leq 0.05$; ##, $p \leq 0.01$ GFP TNF *versus* GFP control (*Ctrl*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

RelB in nonstimulated muscle cells and, to a smaller extent, of p105 and p100 in TNFα-treated cells again suggesting a PGC-1α-dependent increase in the noncanonical NF-κB pathway (Fig. 6*A* and supplemental Fig. S4*A*). In contrast, PGC-1β does not seem to elevate the noncanonical NF-κB family members like PGC-1α. Thus, besides the strong increase in *c*-Rel protein and the more moderate elevation of p105 protein in nonstimulated cells, PGC-1β significantly reduced RelB protein levels (Fig. 6*A* and supplemental Fig. S4*B*). The discrepancy between the PGC-1-mediated repression of *c-Rel* gene expression compared with the elevation of c-Rel protein levels implies post-transcriptional effects in the regulation of this particular NF-κB family member.

Dephosphorylation and Transrepression of p65 Are Potential Molecular Mechanisms for Diminished Cytokine Expression—To further substantiate the findings implying that the PGC-1 coactivators modulate NF-κB activity by preventing p65 phosphorylation, we examined other upstream effectors for this phosphorylation as well as downstream events affecting the DNA-bound NF-κB transcriptional complex. The protein kinase Akt has been implicated in the regulation of the NF-κB signaling pathway upstream of IKKα (36), one of the protein kinases to mediate p65 phosphorylation and of which protein levels are reduced by *Pgc-1α* and *Pgc-1β* overexpression (Fig. 6*A*). As expected, based on these findings, PGC-1α and PGC-1β diminished Akt activation as evident from diminished phospho-Akt (Ser-473) levels normalized to total Akt protein (Fig. 7*A*).

To obtain a more accurate picture of p65 phosphorylation, a time course experiment was performed. It revealed that p65 is phosphorylated at serine 536 after 5 min of TNFα treatment even in the presence of PGC-1α and PGC-1β (Fig. 7, *B* and *C*) implying the possibility that not only an altered kinase profile but also activity of a protein phosphatase might be involved in the PGC-1-mediated modulation of NF-κB phosphorylation. We therefore tested whether pharmacological inhibition of protein phosphatase 2A (PP2A) and PP1, two enzymes that dephosphorylate p65 (37), by okadaic acid abolishes the repression of p65 phosphorylation mediated by PGC-1α and PGC-1β in muscle cells. As expected, okadaic acid powerfully stabilized phosphorylation of p65 at serine 536 (Fig. 7, *D* and *F*). Strikingly, however, *Pgc-1α* and *Pgc-1β* overexpression (Fig. 7, *D* and *E*) still reduced p65 phosphorylation, even in okadaic acid-treated cells (Fig. 7, *D* and *F*). Thus, whereas PP1 and PP2A clearly affect the serine 536 phosphorylation of p65 in our experimental context, these two phosphatases are most likely not involved in the modulation of p65 phosphorylation by the PGC-1s.

Because the phosphorylation status of p65 affects its affinity to cofactors (38), we also determined the expression levels of the coactivator *CBP* (cyclic AMP-responsive element-binding protein-binding protein) and corepressors nuclear receptor corepressor 1 (*Ncor*) and *Smrt*. Of those, only *Smrt* transcript levels were significantly increased by PGC-1α (Fig. 8, *A* and *B*). A well described mechanism of inhibition of inflammatory gene expression is mediated by nuclear receptors and termed tran-

FIGURE 5. **PGC-1s suppress NF-κB transcription activation potential without changing DNA binding or affecting NF-κB expression levels.** A, C2C12 cells were transfected with a wild type and a mutated NF-κB reporter construct, *p65* and increasing amounts of *Pgc-1α* or *Pgc-1β*. Luciferase activity was determined after 24 h and is expressed as ratio of WT reporter to mutated reporter gene expression. ##, $p < 0.01$ PGC-1 *versus* control (*Ctrl*). *B–E*, differentiated C2C12 myotubes overexpressing *Pgc-1α* and *Gfp* (*panels B* and *D*) or *Pgc-1β* and *Gfp* (*panels C* and *E*) were treated with TNFα for 2 h. *B* and *C*, expression of NF-κB family members was determined by real-time PCR. *D* and *E*, DNA binding of NF-κB family members in nuclear extracts was measured by TransAM. Values represent at least 3 independent experiments, mean ± S.E. #, $p \leq 0.05$; ##, $p \leq 0.01$ GFP TNF *versus* GFP control (*Ctrl*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

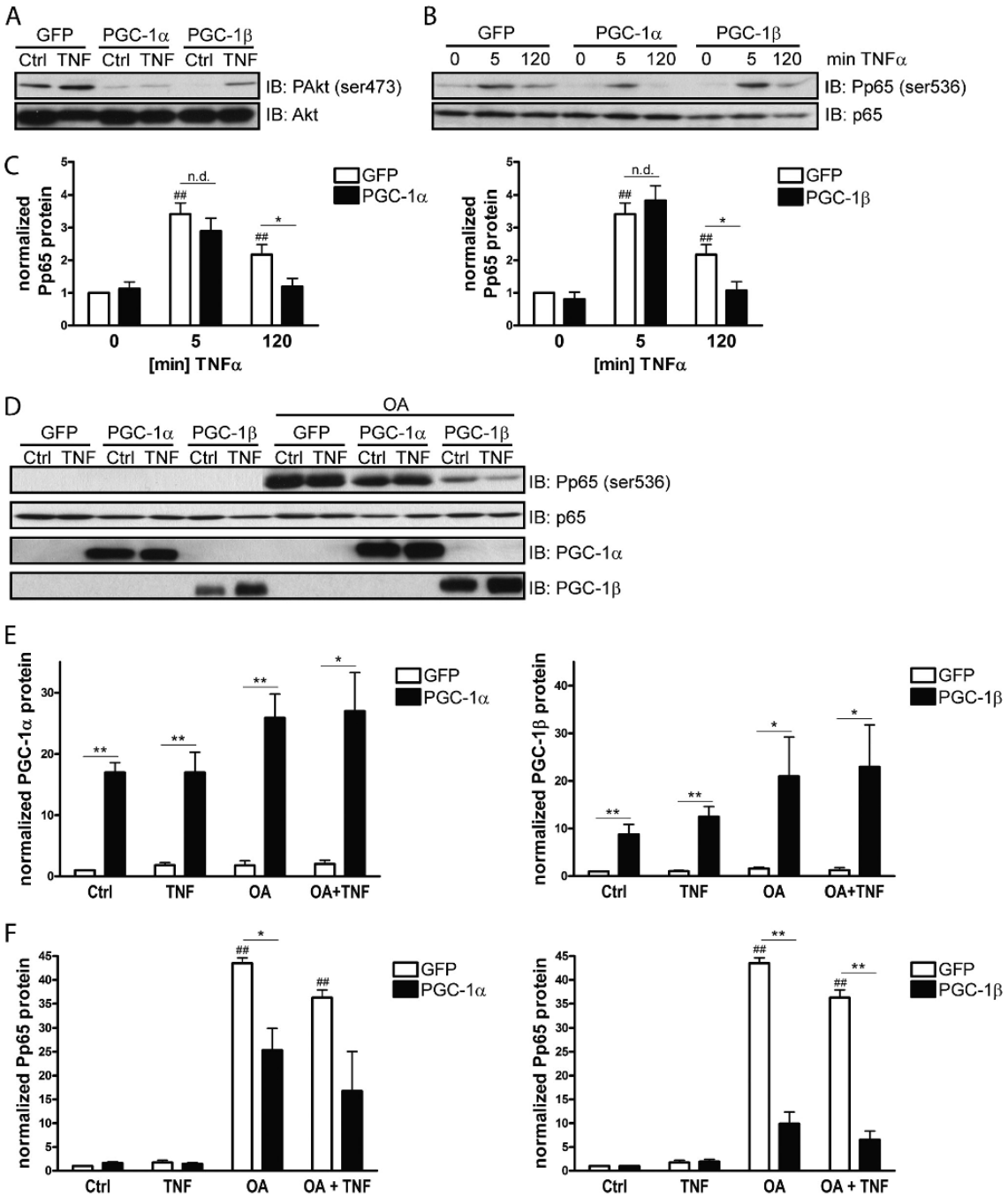


FIGURE 7. **Dephosphorylation of p65 is a potential molecular mechanism for diminished cytokine expression.** *A*, differentiated C2C12 myotubes overexpressing *Pgc-1α*, *Pgc-1β*, and *Gfp* were treated with TNFα for 2 h. Protein levels of Akt and phospho-Akt (Ser-473) were determined by immunoblotting. *B*, differentiated C2C12 myotubes overexpressing *Pgc-1α*, *Pgc-1β*, and *Gfp* were treated with TNFα for 5 min and 2 h, respectively. Protein abundance of phospho-p65 (Ser-536) and total p65 was assessed by immunoblotting (*IB*). *C*, quantification of phospho-p65 (Ser-536) levels from *B*, normalized to α-Tubulin. *D*, differentiated C2C12 myotubes overexpressing *Pgc-1α*, *Pgc-1β*, and *Gfp* were treated with TNFα for 2 h in the presence or absence of okadaic acid (*OA*). Protein abundance of phospho-p65 (Ser-536), total p65, PGC-1α, and PGC-1β was assessed by immunoblotting *E*. *F*, quantification of PGC-1α and PGC-1β protein levels from *D*, normalized to α-Tubulin (*panel E*). Quantification of phospho-p65 (Ser-536) protein levels from *D*, normalized to α-Tubulin (*panel F*). Values represent at least 3 independent experiments, mean ± S.E. #, $p \leq 0.05$; ##, $p \leq 0.01$ GFP OA *versus* GFP control (*Ctrl*); *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP.

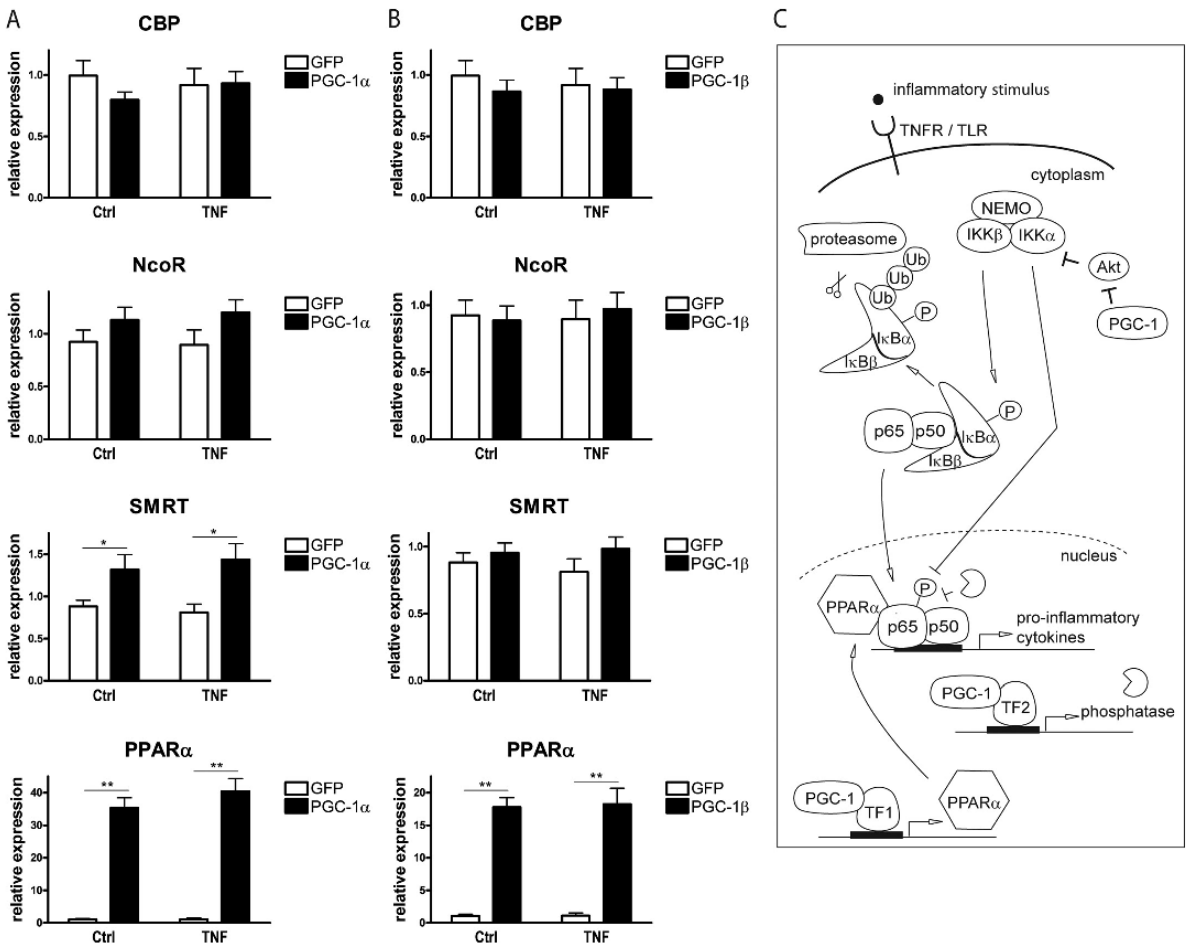


FIGURE 8. **Transrepression of p65 is a potential molecular mechanism for diminished cytokine expression.** *A* and *B*, differentiated C2C12 myotubes overexpressing *Pgc-1α* and *Gfp* (*panel A*) or *Pgc-1β* and *Gfp* (*panel B*) were treated with TNFα for 2 h. Expression of coactivator and corepressor proteins, and nuclear receptor *Pparα* was determined by real-time PCR. *, $p \leq 0.05$; **, $p \leq 0.01$ PGC-1α/β *versus* GFP. *C*, proposed model of PGC-1α/β interference with NF-κB signaling. Gene-, activator- and PGC-1-specific repression of NF-κB target genes is mediated by reduced phosphorylation of p65, transcriptional induction of *Pparα*, and subsequent transrepression as well as activation of an unknown protein phosphatase by both PGC-1s. Furthermore, PGC-1β specifically inhibits *p65* and *p50* expression, whereas PGC-1α elevates members of the NF-κB family involved in the alternative, noncanonical activation (not shown).

srepression (39). We evaluated the gene expression of potential candidates for transrepression and found *Pparα* to be strongly induced by both PGC-1α and PGC-1β in nonstimulated and TNFα-treated muscle cells similar to previous data (40) (Fig. 8, *A* and *B*). PPARα therefore likely contributes to the PGC-1-dependent reduction in NF-κB transcriptional activity by impairing the exchange of corepressors for coactivators that is necessary to effectively initiate transcription. In fact, inhibition of PPARα recovered expression of IL-6 and TNFα in the presence PGC-1β (supplemental Fig. S5).

DISCUSSION

With an aging population and an increasingly sedentary lifestyle, chronic diseases are on the rise. Obesity and its co-morbidities but also some cancers and neurodegeneration have been associated with local and systemic inflammation that worsens disease progression, whereas exercise has beneficial effects in many of these disorders and even acts preventive (41). PGC-1α is a major molecular mediator of exercise in skeletal muscle and its loss not only disturbs metabolic processes but also evokes a local and systemic inflammation (24). In the present report we thus tested the idea that PGC-1 coactivators have anti-inflammatory properties.

Indeed, we confirmed such properties as PGC-1α and PGC-1β were able to diminish the increase in proinflammatory cytokines elicited by different inflammatory stimuli such as TNFα, TLR agonists, and saturated FFAs. We identified the NF-κB pathway as a main target of PGC-1-dependent repression. These results are complemented by *in vivo* findings of Brault and co-workers (42) who showed that NF-κB reporter activity decreases after electroporation of tibialis anterior muscle with *Pgc-1α*/*Pgc-1β* in the context of anti-atrophic effects of PGC-1s. Similar conclusions were derived from experiments in human aortic smooth muscle and endothelial cells where PGC-1α suppressed TNFα-induced *Vcam-1* and *Mcp-1*

expression that contributes to inflammation in atherosclerosis (43).

Mechanistically, PGC-1α and PGC-1β lower phosphorylation of the NF-κB family member p65, which limits its transcriptional activation potential. This is further substantiated by data from muscle-specific PGC-1α transgenic animals that also exhibit reduced p65 phosphorylation (44). Diametrically opposed to this report, another recent publication states that p65 phosphorylation is higher in PGC-1α transgenic muscle before as well as after injection of TNFα (45). The exact role of PGC-1α on muscle inflammation *in vivo* thus remains unresolved. In our cellular model, decreased p65 phosphorylation unequivocally corresponds to the loss of IKKα induction by TNFα in *Pgc-1* overexpressing cells and, interestingly, to diminished Akt activation. Akt-dependent phosphorylation of p65 by IKK is an important mechanism to regulate p65 transactivation potential (36, 46) and thus a good candidate to mediate the reduction observed (Fig. 8*C*). In addition to modulation of kinase activity, the involvement of a phosphatase is suggested by sustained phosphorylation of p65 after TNFα treatment for 5 min even in the presence of PGC-1α and PGC-1β. As okadaic acid treatment did not abrogate differences between conditions, PP2A and PP1 are, however, unlikely to account for this effect. Therefore, further experiments are needed to determine the contribution of different phosphatases in this setting.

The transactivation potential of p65 is controlled by its phosphorylation status as it defines the affinity for cofactors important in suppressing or stimulating transcription of target genes (38). Decreased phosphorylation thus favors interaction with corepressors such as SMRT that decreases p65/p50 transactivation (47). Interestingly, *Smrt* levels are slightly induced by PGC-1α and accordingly could contribute to lower cytokine expression. Stabilization of corepressor complexes on DNA-bound p65 is also fostered by nuclear receptor-mediated transrepression (39). For example, *Pparα* is able to exert anti-inflammatory action by ligand-dependent and -independent transrepressive mechanisms (48, 49). We found a marked induction of PPARα by PGC-1α and PGC-1β, which presumably also contributes to negative regulation of proinflammatory cytokines by transrepression (Fig. 8*C*). This claim is substantiated by the reversal of the repressive PGC-1β effects on proinflammatory cytokine expression when PPARα was inhibited. PPARα regulates lipid metabolism and FFAs are able to serve as ligands for PPARα (50). Thus, the very pronounced suppressive effect of the PGC-1s on inflammatory gene expression observed after FFA treatment might reflect an additional ligand-dependent activation of PPARα leading to an even stronger transrepression in that experimental context.

Besides the reduction in p65 phosphorylation that is exerted by both PGC-1α and PGC-1β, only the latter was further found to repress *p65* and *p50* transcription in the basal state and accordingly the ability of these proteins to bind to DNA response elements. This offers an attractive explanation for the very low cytokine levels observed in the presence of PGC-1β. The NF-κB family members *RelB* and *c-Rel* were also suppressed transcriptionally in the basal state, which puts PGC-1β in the position of a broader anti-inflammatory factor in skeletal muscle. Such anti-inflammatory potential was previously described only in macrophages, where PGC-1β is essential in alternative activation and reactive oxygen species production (27, 28). In contrast, PGC-1α does not alter expression of the classically activated/canonical NF-κB isoforms *p65* and *p50*. However, *Pgc-1α* overexpression induces transcript and protein levels of the alternative isoforms p100/p52 and RelB. This indicates a switch toward noncanonical/alternative NF-κB signaling. A recent publication outlined that alternative signaling via IKKα and RelB induces an oxidative phenotype in muscle driven by PGC-1β (51). Furthermore, activation of *c*-Rel and p50 was suggested to play a role in disuse atrophy (52). Finally, canonical p65 activation was linked to mitochondrial biogenesis in mouse embryonic fibroblasts and liver cells (53, 54). All of these findings together with our data suggest a reciprocal, functional link between NF-κB signaling and oxidative metabolism. A potential induction of the alternative NF-κB pathway by PGC-1α in muscle therefore warrants further investigations.

Strikingly, inflammatory gene transcription was often selectively regulated by the two PGC-1 coactivators depending on the stimulus and specific gene. These findings argue against a general repressive effect of PGC-1α and PGC-1β on tissue inflammation in muscle but rather indicate a specific, fine-tuned effect of these coactivators on NF-κB target genes. For example, whereas TLR1/2-induced cytokine production was suppressed by both PGC-1α and PGC-1β, only PGC-1α was able to block this production after treatment of the cells with a TLR4-specific ligand. Moreover, TLR6/2-induced *Tnfα* expression was also diminished by PGC-1α, which was, however, not the case for TLR6/2-induced expression of *Il-6*. This probably reflects the activation of pathways other than NF-κB downstream of TLRs that are not subject to regulation by PGC-1 or targeted by only one of the PGC-1 isoforms, respectively. Interestingly, whereas neither PGC-1α nor PGC-1β affected expression of TNF receptor 1 (*Tnfr1*) (supplemental Fig. S6, *A* and *B*) both coactivators lowered mRNA levels of *Tlr1*, *Tlr4*, and *Tlr6* (supplemental Fig. 6, *C* and *D*). This down-regulation might presumably contribute to the repressive effects observed with some TLR agonists. The expression pattern of *Tlr1*, *Tlr4*, and *Tlr6* (supplemental Fig. S6, *C* and *D*) further resembles genes in cluster 2 of the microarray, confirming a broader influence of PGC-1s on inflammatory genes. Moreover, the selectivity of the PGC-1s most likely reflects the distinct expression pattern as well as the distinct functions of PGC-1α and PGC-1β in the regulation of skeletal muscle physiology that will have to be further dissected in future experiments. In any case, however, the concept that PGC-1s selectively and specifically repress inflammatory processes and thereby avoid the harmful consequences of broad and general cytokine suppression is compelling. Once understood in greater detail, the therapeutic potential of targeting PGC-1s to modulate specific inflammatory responses in muscle would be immense. Obviously, patients suffering from obesity and type 2 diabetes would profit from lower systemic inflammation mediated by rectified gene expression of *Pgc-1α* and *Pgc-1β* in skeletal muscle. Intriguingly, however, muscle disorders like cachexia, muscular dystrophies, disuse atrophy, or inflammatory myopathy also involve an inflammatory component with activation of the NF-κB pathway (52, 55, 56, 58). In fact, chronic stimulation of

the classical NF-κB pathway in muscle is sufficient to induce muscle wasting (5). Accordingly, mice heterozygous for *p65* or harboring a genetic ablation of the *Ikkβ* gene in the *mdx* background, a model for Duchenne muscular dystrophy, have improved pathology (59). Strikingly, ectopic elevation of PGC-1α in animal models for some of these diseases resulted in an amelioration of fiber damage and muscle functionality, *e.g.* in Duchenne muscular dystrophy, sarcopenia, a mitochondrial myopathy and denervation-induced fiber atrophy (20, 21, 44, 57). Our data now suggest that at least part of the therapeutic effect of PGC-1α in these disease paradigms might stem from the anti-inflammatory effect. It is, thus, tempting to speculate that elevating PGC-1α and/or PGC-1β in muscle would also be beneficial in other conditions of muscle wasting by limiting the detrimental inflammatory component of the disease. In fact, muscle adaptation to endurance training that correlates with increased *Pgc-1α* expression includes an increased resistance against fiber damage, tissue inflammation, and as a consequence, decreased exercise-induced muscle soreness. In contrast to PGC-1α, the implications of the repressive effect of PGC-1β on inflammatory gene expression remain less obvious until the physiological context of PGC-1β regulation and the function of this coactivator in muscle tissue have been more clearly delineated. Nevertheless, by virtue of their ability to reduce NF-κB activation, the PGC-1 coactivators are promising targets to antagonize inflammatory reactions in skeletal muscle associated with a large number of diseases.

Acknowledgments—We thank Dr. Matthias Altmeyer and Dr. Karin Rothgiesser for help with the microarray experiment, Dr. Hubert Rehrauer for advice on statistical analysis of the microarray, and Markus Beer for excellent technical assistance.

REFERENCES

1. Hotamisligil, G. S. (2006) Inflammation and metabolic disorders. *Nature* **444,** 860–867
2. Haffner, S. M. (2006) The metabolic syndrome. Inflammation, diabetes mellitus, and cardiovascular disease. *Am. J. Cardiol.* **97,** 3A–11A
3. Perry, V. H. (2004) The influence of systemic inflammation on inflammation in the brain. Implications for chronic neurodegenerative disease. *Brain Behav. Immun.* **18,** 407–413
4. Arkan, M. C., Hevener, A. L., Greten, F. R., Maeda, S., Li, Z. W., Long, J. M., Wynshaw-Boris, A., Poli, G., Olefsky, J., and Karin, M. (2005) IKK-β links inflammation to obesity-induced insulin resistance. *Nat. Med.* **11,** 191–198
5. Cai, D., Frantz, J. D., Tawa, N. E., Jr., Melendez, P. A., Oh, B. C., Lidov, H. G., Hasselgren, P. O., Frontera, W. R., Lee, J., Glass, D. J., and Shoelson, S. E. (2004) IKKβ/NF-κB activation causes severe muscle wasting in mice. *Cell* **119,** 285–298
6. Zhang, X., Zhang, G., Zhang, H., Karin, M., Bai, H., and Cai, D. (2008) Hypothalamic IKKβ/NF-κB and ER stress link overnutrition to energy imbalance and obesity. *Cell* **135,** 61–73
7. Shi, H., Kokoeva, M. V., Inouye, K., Tzameli, I., Yin, H., and Flier, J. S. (2006) TLR4 links innate immunity and fatty acid-induced insulin resistance. *J. Clin. Invest.* **116,** 3015–3025
8. Ghosh, S., and Hayden, M. S. (2008) New regulators of NF-κB in inflammation. *Nat. Rev. Immunol.* **8,** 837–848
9. Sakurai, H., Chiba, H., Miyoshi, H., Sugita, T., and Toriumi, W. (1999) IκB kinases phosphorylate NF-κB p65 subunit on serine 536 in the transactivation domain. *J. Biol. Chem.* **274,** 30353–30356
10. Buss, H., Dörrie, A., Schmitz, M. L., Hoffmann, E., Resch, K., and Kracht, M. (2004) Constitutive and interleukin-1-inducible phosphorylation of p65 NF-κB at serine 536 is mediated by multiple protein kinases including IκB kinase (IKK)-α, IKKβ, IKKε, TRAF family member-associated (TANK)-binding kinase 1 (TBK1), and an unknown kinase and couples p65 to TATA-binding protein-associated factor II31-mediated interleukin-8 transcription. *J. Biol. Chem.* **279,** 55633–55643
11. Gleeson, M. (2007) Immune function in sport and exercise. *J. Appl. Physiol.* **103,** 693–699
12. Knowler, W. C., Barrett-Connor, E., Fowler, S. E., Hamman, R. F., Lachin, J. M., Walker, E. A., and Nathan, D. M. (2002) Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N. Engl. J. Med.* **346,** 393–403
13. Fiatarone, M. A., O'Neill, E. F., Ryan, N. D., Clements, K. M., Solares, G. R., Nelson, M. E., Roberts, S. B., Kehayias, J. J., Lipsitz, L. A., and Evans, W. J. (1994) Exercise training and nutritional supplementation for physical frailty in very elderly people. *N. Engl. J. Med.* **330,** 1769–1775
14. Tillerson, J. L., Caudle, W. M., Reverón, M. E., and Miller, G. W. (2003) Exercise induces behavioral recovery and attenuates neurochemical deficits in rodent models of Parkinson's disease. *Neuroscience* **119,** 899–911
15. Puigserver, P., Wu, Z., Park, C. W., Graves, R., Wright, M., and Spiegelman, B. M. (1998) A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. *Cell* **92,** 829–839
16. Pilegaard, H., Saltin, B., and Neufer, P. D. (2003) Exercise induces transient transcriptional activation of the PGC-1α gene in human skeletal muscle. *J. Physiol.* **546,** 851–858
17. Puigserver, P., Adelmant, G., Wu, Z., Fan, M., Xu, J., O'Malley, B., and Spiegelman, B. M. (1999) Activation of PPARγ coactivator-1 through transcription factor docking. *Science* **286,** 1368–1371
18. Lin, J., Wu, H., Tarr, P. T., Zhang, C. Y., Wu, Z., Boss, O., Michael, L. F., Puigserver, P., Isotani, E., Olson, E. N., Lowell, B. B., Bassel-Duby, R., and Spiegelman, B. M. (2002) Transcriptional co-activator PGC-1α drives the formation of slow-twitch muscle fibres. *Nature* **418,** 797–801
19. Calvo, J. A., Daniels, T. G., Wang, X., Paul, A., Lin, J., Spiegelman, B. M., Stevenson, S. C., and Rangwala, S. M. (2008) Muscle-specific expression of PPARγ coactivator-1α improves exercise performance and increases peak oxygen uptake. *J. Appl. Physiol.* **104,** 1304–1312
20. Handschin, C., Kobayashi, Y. M., Chin, S., Seale, P., Campbell, K. P., and Spiegelman, B. M. (2007) PGC-1α regulates the neuromuscular junction program and ameliorates Duchenne muscular dystrophy. *Genes Dev.* **21,** 770–783
21. Sandri, M., Lin, J., Handschin, C., Yang, W., Arany, Z. P., Lecker, S. H., Goldberg, A. L., and Spiegelman, B. M. (2006) PGC-1α protects skeletal muscle from atrophy by suppressing FoxO3 action and atrophy-specific gene transcription. *Proc. Natl. Acad. Sci. U.S.A.* **103,** 16260–16265
22. Hanai, J., Cao, P., Tanksale, P., Imamura, S., Koshimizu, E., Zhao, J., Kishi, S., Yamashita, M., Phillips, P. S., Sukhatme, V. P., and Lecker, S. H. (2007) The muscle-specific ubiquitin ligase atrogin-1/MAFbx mediates statin-induced muscle toxicity. *J. Clin. Invest.* **117,** 3940–3951
23. Handschin, C., Chin, S., Li, P., Liu, F., Maratos-Flier, E., Lebrasseur, N. K., Yan, Z., and Spiegelman, B. M. (2007) Skeletal muscle fiber-type switching, exercise intolerance, and myopathy in PGC-1α muscle-specific knock-out animals. *J. Biol. Chem.* **282,** 30014–30021
24. Handschin, C., Choi, C. S., Chin, S., Kim, S., Kawamori, D., Kurpad, A. J., Neubauer, N., Hu, J., Mootha, V. K., Kim, Y. B., Kulkarni, R. N., Shulman, G. I., and Spiegelman, B. M. (2007) Abnormal glucose homeostasis in skeletal muscle-specific PGC-1α knockout mice reveals skeletal muscle-pancreatic β cell crosstalk. *J. Clin. Invest.* **117,** 3463–3474
25. Patti, M. E., Butte, A. J., Crunkhorn, S., Cusi, K., Berria, R., Kashyap, S., Miyazaki, Y., Kohane, I., Costello, M., Saccone, R., Landaker, E. J., Goldfine, A. B., Mun, E., DeFronzo, R., Finlayson, J., Kahn, C. R., and Mandarino, L. J. (2003) Coordinated reduction of genes of oxidative metabolism in humans with insulin resistance and diabetes. Potential role of PGC1 and NRF1. *Proc. Natl. Acad. Sci. U.S.A.* **100,** 8466–8471
26. Arany, Z., Lebrasseur, N., Morris, C., Smith, E., Yang, W., Ma, Y., Chin, S., and Spiegelman, B. M. (2007) The transcriptional coactivator PGC-1β drives the formation of oxidative type IIX fibers in skeletal muscle. *Cell Metab.* **5,** 35–46
27. Vats, D., Mukundan, L., Odegaard, J. I., Zhang, L., Smith, K. L., Morel,

C. R., Wagner, R. A., Greaves, D. R., Murray, P. J., and Chawla, A. (2006) Oxidative metabolism and PGC-1β attenuate macrophage-mediated inflammation. *Cell Metab.* **4,** 13–24
28. Sonoda, J., Laganière, J., Mehl, I. R., Barish, G. D., Chong, L. W., Li, X., Scheffler, I. E., Mock, D. C., Bataille, A. R., Robert, F., Lee, C. H., Giguère, V., and Evans, R. M. (2007) Nuclear receptor ERRα and coactivator PGC-1β are effectors of IFN-γ-induced host defense. *Genes Dev.* **21,** 1909–1920
29. Jayne, S., Rothgiesser, K. M., and Hottiger, M. O. (2009) CARM1 but not its enzymatic activity is required for transcriptional coactivation of NF-κB-dependent gene expression. *J. Mol. Biol.* **394,** 485–495
30. Al-Shahrour, F., Díaz-Uriarte, R., and Dopazo, J. (2004) FatiGO. A web tool for finding significant associations of Gene Ontology terms with groups of genes. *Bioinformatics* **20,** 578–580
31. Arnold, P., Erb, I., Pachkov, M., Molina, N., and van Nimwegen, E. (2012) MotEvo. Integrated Bayesian probabilistic methods for inferring regulatory sites and motifs on multiple alignments of DNA sequences. *Bioinformatics* **28,** 487–494
32. FANTOM Consortium, Suzuki, H., Forrest, A. R., van Nimwegen, E., Daub, C. O., Balwierz, P. J., Irvine, K. M., Lassmann, T., Ravasi, T., Hasegawa, Y., de Hoon, M. J., *et al.* (2009) The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. *Nat. Genet.* **41,** 553–562
33. Coll, T., Eyre, E., Rodríguez-Calvo, R., Palomer, X., Sánchez, R. M., Merlos, M., Laguna, J. C., and Vázquez-Carrera, M. (2008) Oleate reverses palmitate-induced insulin resistance and inflammation in skeletal muscle cells. *J. Biol. Chem.* **283,** 11107–11116
34. Arany, Z., Foo, S. Y., Ma, Y., Ruas, J. L., Bommi-Reddy, A., Girnun, G., Cooper, M., Laznik, D., Chinsomboon, J., Rangwala, S. M., Baek, K. H., Rosenzweig, A., and Spiegelman, B. M. (2008) HIF-independent regulation of VEGF and angiogenesis by the transcriptional coactivator PGC-1α. *Nature* **451,** 1008–1012
35. Bhatnagar, S., Panguluri, S. K., Gupta, S. K., Dahiya, S., Lundy, R. F., and Kumar, A. (2010) Tumor necrosis factor-α regulates distinct molecular pathways and gene networks in cultured skeletal muscle cells. *PloS One* **5,** e13262
36. Sizemore, N., Lerner, N., Dombrowski, N., Sakurai, H., and Stark, G. R. (2002) Distinct roles of the IκB kinase α and β subunits in liberating nuclear factor κB (NF-κB) from IκB and in phosphorylating the p65 subunit of NF-κB. *J. Biol. Chem.* **277,** 3863–3869
37. Yang, J., Fan, G. H., Wadzinski, B. E., Sakurai, H., and Richmond, A. (2001) Protein phosphatase 2A interacts with and directly dephosphorylates RelA. *J. Biol. Chem.* **276,** 47828–47833
38. Zhong, H., Voll, R. E., and Ghosh, S. (1998) Phosphorylation of NF-κB p65 by PKA stimulates transcriptional activity by promoting a novel bivalent interaction with the coactivator CBP/p300. *Mol. Cell* **1,** 661–671
39. Ghisletti, S., Huang, W., Ogawa, S., Pascual, G., Lin, M. E., Willson, T. M., Rosenfeld, M. G., and Glass, C. K. (2007) Parallel SUMOylation-dependent pathways mediate gene- and signal-specific transrepression by LXRs and PPARγ. *Mol. cell* **25,** 57–70
40. Huss, J. M., Torra, I. P., Staels, B., Giguère, V., and Kelly, D. P. (2004) Estrogen-related receptor α directs peroxisome proliferator-activated receptor α signaling in the transcriptional control of energy metabolism in cardiac and skeletal muscle. *Mol. Cell. Biol.* **24,** 9079–9091
41. Handschin, C., and Spiegelman, B. M. (2008) The role of exercise and PGC1α in inflammation and chronic disease. *Nature* **454,** 463–469
42. Brault, J. J., Jespersen, J. G., and Goldberg, A. L. (2010) Peroxisome proliferator-activated receptor γ coactivator 1α or 1β overexpression inhibits muscle protein degradation, induction of ubiquitin ligases, and disuse atrophy. *J. Biol. Chem.* **285,** 19460–19471
43. Kim, H. J., Park, K. G., Yoo, E. K., Kim, Y. H., Kim, Y. N., Kim, H. S., Kim, H. T., Park, J. Y., Lee, K. U., Jang, W. G., Kim, J. G, Kim, B. W., and Lee, I. K. (2007) Effects of PGC-1α on TNF-α-induced MCP-1 and VCAM-1 expression and NF-κB activation in human aortic smooth muscle and endothelial cells. *Antioxid. Redox Signal.* **9,** 301–307
44. Wenz, T., Rossi, S. G., Rotundo, R. L., Spiegelman, B. M., and Moraes, C. T. (2009) Increased muscle PGC-1α expression protects from sarcopenia and metabolic disease during aging. *Proc. Natl. Acad. Sci. U.S.A.* **106,** 20405–20410
45. Olesen, J., Larsson, S., Iversen, N., Yousafzai, S., Hellsten, Y., and Pilegaard, H. (2012) Skeletal muscle PGC-1α is required for maintaining an acute LPS-induced TNFα response. *PloS One* **7,** e32222
46. Madrid, L. V., Mayo, M. W., Reuther, J. Y., and Baldwin, A. S., Jr. (2001) Akt stimulates the transactivation potential of the RelA/p65 Subunit of NF-κB through utilization of the IκB kinase and activation of the mitogen-activated protein kinase p38. *J. Biol. Chem.* **276,** 18934–18940
47. Lee, S. K., Kim, J. H., Lee, Y. C., Cheong, J., and Lee, J. W. (2000) Silencing mediator of retinoic acid and thyroid hormone receptors, as a novel transcriptional corepressor molecule of activating protein-1, nuclear factor-κB, and serum response factor. *J. Biol. Chem.* **275,** 12470–12474
48. Staels, B., Koenig, W., Habib, A., Merval, R., Lebret, M., Torra, I. P., Delerive, P., Fadel, A., Chinetti, G., Fruchart, J. C., Najib, J., Maclouf, J., and Tedgui, A. (1998) Activation of human aortic smooth-muscle cells is inhibited by PPARα but not by PPARγ activators. *Nature* **393,** 790–793
49. Blanquart, C., Mansouri, R., Paumelle, R., Fruchart, J. C., Staels, B., and Glineur, C. (2004) The protein kinase C signaling pathway regulates a molecular switch between transactivation and transrepression activity of the peroxisome proliferator-activated receptor α. *Mol. Endocrinol.* **18,** 1906–1918
50. Kliewer, S. A., Sundseth, S. S., Jones, S. A., Brown, P. J., Wisely, G. B., Koble, C. S., Devchand, P., Wahli, W., Willson, T. M., Lenhard, J. M., and Lehmann, J. M. (1997) Fatty acids and eicosanoids regulate gene expression through direct interactions with peroxisome proliferator-activated receptors α and γ. *Proc. Natl. Acad. Sci. U.S.A.* **94,** 4318–4323
51. Bakkar, N., Ladner, K., Canan, B. D., Liyanarachchi, S., Bal, N. C., Pant, M., Periasamy, M., Li, Q., Janssen, P. M., and Guttridge, D. C. (2012) IKKα and alternative NF-κB regulate PGC-1β to promote oxidative muscle metabolism. *J. Cell Biol.* **196,** 497–511
52. Hunter, R. B., Stevenson, E., Koncarevic, A., Mitchell-Felton, H., Essig, D. A., and Kandarian, S. C. (2002) Activation of an alternative NF-κB pathway in skeletal muscle during disuse atrophy. *FASEB J.* **16,** 529–538
53. Mauro, C., Leow, S. C., Anso, E., Rocha, S., Thotakura, A. K., Tornatore, L., Moretti, M., De Smaele, E., Beg, A. A., Tergaonkar, V., Chandel, N. S., and Franzoso, G. (2011) NF-κB controls energy homeostasis and metabolic adaptation by up-regulating mitochondrial respiration. *Nat. Cell. Biol.* **13,** 1272–1279
54. Suliman, H. B., Sweeney, T. E., Withers, C. M., and Piantadosi, C. A. (2010) Co-regulation of nuclear respiratory factor-1 by NFκB and CREB links LPS-induced inflammation to mitochondrial biogenesis. *J. Cell Sci.* **123,** 2565–2575
55. Tracey, K. J., Wei, H., Manogue, K. R., Fong, Y., Hesse, D. G., Nguyen, H. T., Kuo, G. C., Beutler, B., Cotran, R. S., and Cerami, A. (1988) Cachectin/tumor necrosis factor induces cachexia, anemia, and inflammation. *J. Exp. Med.* **167,** 1211–1227
56. Spencer, M. J., Walsh, C. M., Dorshkind, K. A., Rodriguez, E. M., and Tidball, J. G. (1997) Myonuclear apoptosis in dystrophic mdx muscle occurs by perforin-mediated cytotoxicity. *J. Clin. Invest.* **99,** 2745–2751
57. Wenz, T., Diaz, F., Spiegelman, B. M., and Moraes, C. T. (2008) Activation of the PPAR/PGC-1α pathway prevents a bioenergetic deficit and effectively improves a mitochondrial myopathy phenotype. *Cell Metab.* **8,** 249–256
58. Creus, K. K., De Paepe, B., Werbrouck, B. F., Vervaet, V., Weis, J., and De Bleecker, J. L. (2009) Distribution of the NF-κB complex in the inflammatory exudates characterizing the idiopathic inflammatory myopathies. *Ann. N.Y. Acad. Sci.* **1173,** 370–377
59. Acharyya, S., Villalta, S. A., Bakkar, N., Bupha-Intr, T., Janssen, P. M., Carathers, M., Li, Z. W., Beg, A. A., Ghosh, S., Sahenk, Z., Weinstein, M., Gardner, K. L., Rafael-Fortney, J. A., Karin, M., Tidball, J. G., Baldwin, A. S., and Guttridge, D. C. (2007) Interplay of IKK/NF-κB signaling in macrophages and myofibers promotes muscle degeneration in Duchenne muscular dystrophy. *J. Clin. Invest.* **117,** 889–901

Appendix 2

MCB
Journals.ASM.org

The Corepressor NCoR1 Antagonizes PGC-1α and Estrogen-Related Receptor α in the Regulation of Skeletal Muscle Function and Oxidative Metabolism

Joaquín Pérez-Schindler,[a] Serge Summermatter,[a] Silvia Salatino,[a] Francesco Zorzato,[b,c] Markus Beer,[a] Piotr J. Balwierz,[a,d] Erik van Nimwegen,[a,d] Jérôme N. Feige,[e,f] Johan Auwerx,[f] and Christoph Handschin[a]

Biozentrum, University of Basel, Basel, Switzerland[a]; Departments of Anesthesia and Biomedicine, Basel University Hospital, Basel, Switzerland[b]; Department of Experimental and Diagnostic Medicine, University of Ferrara, Ferrara, Italy[c]; Swiss Institute of Bioinformatics, Basel, Switzerland[d]; Novartis Institute for Biomedical Research, Basel, Switzerland[e]; and Laboratory for Integrative and Systems Physiology, Ecole Polytechnique Fédérale de Lausanne, Lausanne, Switzerland[f]

Skeletal muscle exhibits a high plasticity and accordingly can quickly adapt to different physiological and pathological stimuli by changing its phenotype largely through diverse epigenetic mechanisms. The nuclear receptor corepressor 1 (NCoR1) has the ability to mediate gene repression; however, its role in regulating biological programs in skeletal muscle is still poorly understood. We therefore studied the mechanistic and functional aspects of NCoR1 function in this tissue. NCoR1 muscle-specific knockout mice exhibited a 7.2% higher peak oxygen consumption (VO_{2peak}), a 11% reduction in maximal isometric force, and increased *ex vivo* fatigue resistance during maximal stimulation. Interestingly, global gene expression analysis revealed a high overlap between the effects of NCoR1 deletion and peroxisome proliferator-activated receptor gamma (PPARγ) coactivator 1α (PGC-1α) overexpression on oxidative metabolism in muscle. Importantly, PPARβ/δ and estrogen-related receptor α (ERRα) were identified as common targets of NCoR1 and PGC-1α with opposing effects on the transcriptional activity of these nuclear receptors. In fact, the repressive effect of NCoR1 on oxidative phosphorylation gene expression specifically antagonizes PGC-1α-mediated coactivation of ERRα. We therefore delineated the molecular mechanism by which a transcriptional network controlled by corepressor and coactivator proteins determines the metabolic properties of skeletal muscle, thus representing a potential therapeutic target for metabolic diseases.

Improved muscle performance is directly linked to a lower prevalence of metabolic diseases (9, 50). In fact, while physical exercise and training can lower morbidity and mortality, physical inactivity has been recognized as one of the main risk factors for these pathologies (8). Lower whole-body aerobic capacity, muscle mitochondrial content, and oxidative activity, which all correlate with a sedentary lifestyle, contribute to the development of metabolic disorders (9, 25, 34, 38). Therefore, maintenance or improvement of skeletal muscle function, especially its oxidative metabolism, should be considered among the first interventions in the treatment and prevention of metabolic diseases.

Skeletal muscle is a highly plastic tissue that can quickly adapt to different physiological (e.g., exercise) and pathological (e.g., overnutrition) stimuli. In fact, muscle fibers can change their gene expression profile and phenotype to a great extent through diverse epigenetic mechanisms (3, 6, 31). Accordingly, muscle remodeling is highly regulated by different transcription factors and coregulator complexes, which are able to modify chromatin structure and thereby regulate gene transcription (27, 41). The nuclear receptor corepressor 1 (NCoR1) is a ubiquitously expressed corepressor, originally identified as the mediator of ligand-independent transcriptional repression of the thyroid hormone and the retinoic acid receptor (22). NCoR1 interacts with several transcription factors through its receptor interaction domains located in the C terminus (48). However, because NCoR1 lacks intrinsic histone deacetylase (HDAC) activity, it regulates gene transcription by forming a large protein complex in which G protein pathway suppressor 2 (GPS2), transducin β-like 1 (TBL1), TBL-related 1 (TBLR1), and HDAC3 represent the core subunits (52). In fact, the NCoR1-HDAC3 interaction plays an essential role in the control of gene transcription, since HDAC3 is directly activated by the deacetylase activation domain (DAD) of NCoR1 (23).

NCoR1 interacts with different proteins that play an important role in muscle physiology, such as peroxisome proliferator-activated receptors (PPAR) and p85α (15, 32), although its role in skeletal muscle remains largely enigmatic. Cell culture experiments implied that NCoR1 modulates myoblast differentiation through the regulation of the expression and transcriptional activity of several transcription factors, e.g., MyoD, TRα1, and Csl (5, 10, 26). The role of NCoR1 *in vivo* is not well understood because *Ncor1*$^{-/-}$ mice embryos die during gestation (24). Recently, conditional knockout models revealed that NCoR1 is an important player in skeletal muscle and adipose tissue energy metabolism (28, 51). In skeletal muscle, NCoR1 deletion enhances oxidative metabolism and slightly improves insulin tolerance under a high-fat diet (51). Specific genetic ablation of NCoR1 in white adipose tissue lowers inflammation and improves whole-body insulin sensitivity (28). However, the mechanism by which NCoR1 deletion results in these effects is not well understood.

Received 28 June 2012 Returned for modification 16 July 2012
Accepted 23 September 2012
Published ahead of print 1 October 2012
Address correspondence to Christoph Handschin, christoph.handschin@unibas.ch.
Supplemental material for this article may be found at http://mcb.asm.org/.

doi:10.1128/MCB.00877-12

NCoR1 is expressed in adult glycolytic and oxidative muscles at equal levels (42, 43). Chronic low-frequency stimulation of rat hind limb and acute endurance exercise in mice repress the expression of NCoR1 in this tissue (42, 51). Interestingly, by using knockout mice, it has been demonstrated that the formation of slow oxidative muscle fibers is negatively regulated by class IIa HDACs (39), which depend on interaction with the NCoR1-HDAC3 complex in order to induce protein deacetylation (14). Consistent with this, the global disruption of the NCoR1-HDAC3 complex in mice results in improved energy metabolism (e.g., higher insulin sensitivity and oxygen consumption) and altered circadian behavior (1). The aim of this study was to further investigate the role of NCoR1 in different aspects of skeletal muscle function (e.g., force generation) and, importantly, to elucidate the elusive mechanism by which NCoR1 regulates oxidative metabolism in this tissue.

MATERIALS AND METHODS

Animal housing and NCoR1 MKO mouse generation. Mice were housed in a conventional facility with a 12-h night/12-h day cycle with free access to food and water. All experiments were performed on adult male mice with the approval of the Swiss authorities. NCoR1 muscle-specific knockout (MKO) animals were generated as previously described (51). Briefly, *Ncor1*$^{loxP/loxP}$ mice were crossed with *HSA-Cre* transgenic mice to generate NCoR1 MKO mice. *Ncor1*$^{loxP/loxP}$ animals without *Cre* expression were used as control (CON) mice. No overt phenotypic differences between CON and wild-type (WT) mice were observed. Genotyping was performed from tail biopsy specimens by PCR using specific primer pairs to detect the presence of the 5′ and 3′ *loxP* sites. The presence of the 5′ *loxP* site resulted in an amplicon of 450 bp (WT allele, 403 bp), while the presence of the 3′ *loxP* site resulted in an amplicon of 346 bp (WT allele, 207 bp) (see Fig. S1A in the supplemental material). Specific primer pairs to detect *Cre* recombinase resulted in an amplicon of 320 bp in NCoR1 MKO animals (see Fig. S1A). In addition, using muscle samples, recombination was confirmed by PCR using the forward and reverse primers used to detect the 5′ and 3′ *loxP* sites, respectively. Consequently, a 246-bp band was detected exclusively in NCoR1 MKO animals (see Fig. S1B). The recombination of the *Ncor1* floxed allele decreased its mRNA specifically in skeletal and, to a lesser extent, cardiac muscle compared to that in CON mice (see Fig. S1C). Importantly, previous work has indicated that the slight decrease of NCoR1 mRNA in the heart of MKO mice does not affect cardiac morphology and function (51).

The generation and characterization of transgenic mice expressing peroxisome proliferator-activated receptor gamma (PPARγ) coactivator 1α (PGC-1α) under the control of the MCK promoter has been published (30). These PGC-1α muscle-specific transgenic (mTg) mice exhibit a 9.5-fold increase in skeletal muscle PGC-1α mRNA (see Fig. S3A in the supplemental material), higher oxidative metabolism, and improved exercise performance (11, 30).

Exercise performance assessment. Animals were acclimatized to treadmill running for 2 days. On the first day, mice ran in a closed treadmill (Columbus Instruments) for 5 min at 10 m/min with a 0° slope, followed by 5 min at 14 m/min with a 5° incline. On the second day, animals ran for 5 min at 10 m/min with a 5° incline, followed by 5 min at 14 m/min. To determine maximal exercise performance and peak oxygen consumption (VO_{2peak}), indirect calorimetry was performed during a maximal exercise test. Thus, 2 days after the acclimatization, mice were placed in a closed treadmill for 6 min at 0 m/min with a 5° incline. Subsequently, the test started at 8 m/min for 3 min with a 5° incline and the speed was increased 2 m/min every 3 min until exhaustion.

To determine endurance performance, mice were placed in an open treadmill (Columbus Instruments) for 5 min at 0 m/min with a 5° incline, followed by 5 min at 8 m/min. Mice ran for 20 min at 60, 70, 80, and 90% of the maximal speed reached in the maximal exercise test (average of the group), and then the speed increased to 100% of the maximal speed until exhaustion. The endurance test was performed at least 3 days after the maximal test.

Blood lactate analysis. Blood lactate was measured from the tail vein of mice fasted overnight (16 h) or fed animals before and after different time points after treadmill running (see the maximal exercise test) using a lactate meter (Nova Biomedical).

***In vivo* measurement of muscle contractility.** Grip strength of the fore and hind limbs was measured with a grip strength meter (Chatillon). To determine the maximal strength, three measurements were performed with at least 60 s of recovery between each repetition, and the maximum value obtained was used for the analysis. To assess isometric fatigue resistance, mice were placed on top of an elevated grid, and the maximum time that they could remain on the inverted grid was recorded.

***Ex vivo* determination of muscle contractility.** Maximal force and fatigue resistance of extensor digitorum longus (EDL) and soleus were measured with a muscle testing setup (Heidelberg Scientific Instruments). Contraction amplitude was digitalized at 4 kHz with an AD Instruments converter. After the determination of the optimal length, force generation of EDL and soleus during a single twitch was measured in response to a 15-V pulse for 0.5 ms. Tetanic contraction was assessed in response to a 15-V pulse at 150 Hz for 400 ms for EDL and 1,100 ms for soleus. Maximal force and kinetics of the single twitch and tetanus were recorded and analyzed.

After 10 min of recovery following the tetanic stimulation, fatigue resistance of EDL and soleus was assessed under two different protocols. First, a long-interval protocol was performed by stimulating the muscles with a 15-V pulse at 150 Hz for 350 ms at 3.6-s intervals during 6 min for EDL and 10 min for soleus, followed by 10 min of recovery. Subsequently, a short-interval protocol was performed by stimulating the muscles with a 15-V pulse at 150 Hz for 350 ms at 1-s intervals during 2 min for EDL and 3 min for soleus. Changes in force generation are expressed as a percentage of initial force (first tetanus) (18).

Cell culture experiments. C_2C_{12} myoblasts were grown in Dulbecco's modified Eagle's medium (DMEM) supplemented with 10% fetal bovine serum and 1% penicillin-streptomycin (growth medium). To induce differentiation, growth medium of 90 to 95% confluent myoblasts was changed to DMEM supplemented with 2% horse serum and 1% penicillin-streptomycin (differentiation medium). Cells were maintained at 37°C, 95% O_2, and 5% CO_2.

Experiments using C_2C_{12} cells were performed on fully differentiated myotubes 4 days after differentiation was induced. For knockdown experiments, different adenoviruses containing specific short hairpin RNAs (shRNAs) against NCoR1, PGC-1α, PGC-1β, and LacZ (control) were used. Cells were incubated with the corresponding adenovirus for 24 h, and then adenovirus-containing medium was exchanged for fresh differentiation medium for another 24 h. In addition, cells were treated for 48 h with 0.2% dimethylsulfoxide (DMSO) (as a control), 10 μM XCT790 (Sigma-Aldrich) to inhibit estrogen-related receptor α (ERRα), or 1 μM GSK0660 (Sigma-Aldrich) to inhibit PPARβ/δ, together with the corresponding adenovirus. Three independent experiments each were performed in triplicate.

Luciferase assays were performed on 12-well plates using COS-7 cells (African green monkey kidney fibroblast-like cells) grown in growth medium without antibiotics. COS-7 cells were used in our experiments as a heterologous experimental system because of their high transfection efficiency and suitability for expressing constructs containing simian virus 40 (SV40) promoters. Cells were transfected using Lipofectamine 2000 (Invitrogen) with 0.1 μg pRL-SV40 (E2231; Promega), 0.3 μg pPPRE X3-TK-luc (1015; Bruce Spiegelman; Addgene), 0.3 μg pERRE-luc (gift of Junichi Sadoshima) (36), 0.4 μg pBABE puro PPAR delta (8891; Bruce Spiegelman; Addgene), 0.4 μg pERRα (gift of Vincent Giguère), 0.4 μg pFlag-NCoR (gift of Christopher K. Glass), and 0.4 μg pAd-Track HA PGC-1α (14427; Pere Puigserver; Addgene). The total amount of plasmid DNA was kept constant at 1.6 μg per well by using the control plasmid

pAdtrack-CMV (ATCC). Twenty-four hours after transfection, cells were lysed with 250 μl of 1× passive lysis buffer (Promega), and luciferase activity was measured in 75 μl of lysate in a 96-well plate using the Dual-Glo luciferase assay system (Promega). *Renilla* (pRL-SV40) luciferase activity was used for normalization. Five independent experiments each were performed in triplicate.

RNA isolation and real-time PCR. Total RNA was isolated from C_2C_{12} myotubes, liver, kidney, white adipose tissue, heart, gastrocnemius (GAS), soleus, plantaris, tibialis anterior, extensor digitorum longus, and quadriceps from NCoR1 MKO mice or gastrocnemius from PGC-1α mTg animals using lysing matrix tubes (MP Biomedicals) and TRI reagent (Sigma-Aldrich) according to the manufacturer's instructions. RNA concentration was measured with a NanoDrop 1000 spectrophotometer (Thermo Scientific). One microgram of RNA was treated with DNase I (Invitrogen) and then reversed transcribed using hexanucleotide mix (Roche) and SuperScript II reverse transcriptase (Invitrogen). The level of relative mRNA was quantified by real-time PCR on a StepOnePlus system (Applied Biosystems) using Power SYBR green PCR master mix (Applied Biosystems). The analysis of the mRNA was performed by the $\Delta\Delta C_T$ method using TATA binding protein (TBP) as the endogenous control. Primers used for target genes and TBP had the same PCR efficiency. TBP transcript levels were not different between genotypes or different experimental conditions. Primer sequences can be found in Table S3 in the supplemental material.

Mitochondrial DNA measurement. DNA was isolated from gastrocnemius using a NucleoSpin tissue kit (Macherey-Nagel). Real-time PCR analysis was performed to measure COX2 (mitochondrial DNA) and β-globin (nuclear DNA) levels.

Skeletal muscle staining. Oxidative fibers were detected by NADH staining of 12-μm cross-sections from the tibialis anterior. Staining was performed by exposing the sections to 0.8 mg/ml NADH in the presence of 1 mg/ml nitroblue tetrazolium. Periodic acid-Schiff (PAS) staining was performed with a PAS kit by following the manufacturer's instructions (Sigma-Aldrich).

Protein isolation. Tissue samples were powdered on dry ice and homogenized with a polytron device in 300 μl of ice-cold lysis buffer (50 mM Tris-HCl, pH 7.5, 250 mM sucrose, 1 mM EDTA, 1 mM EGTA, 0.25% NP-40 substitute, 50 mM NaF, 5 mM $Na_4P_2O_7$, 0.1% dithiothreitol [DTT], and fresh protease and phosphatase inhibitor cocktail). Samples then were shaken at 1,300 rpm for 30 min at 4°C. Samples were subsequently centrifuged at 13,000 × *g* for 10 min at 4°C, and the protein concentration of the supernatant was determined by the Bradford assay (Bio-Rad). Equal aliquots of protein were boiled for 5 min in Laemmli sample buffer (250 mM Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 0.01% bromophenol blue, and 5% β-mercaptoethanol).

Western blotting. Samples were separated on SDS-polyacrylamide gels and then transferred to nitrocellulose membranes for 60 min. Membranes were blocked for 1 h in 3% milk, Tris-buffered saline, and 0.1% Tween 20 (TBST) before overnight incubation at 4°C with the appropriate primary antibody in TBST (1:1,000 dilution). Proteins were detected with a primary antibody to p-AMPKα^{T172} (2535; Cell Signaling) and AMPKα (2603; Cell Signaling). As a loading control, eEF2 (2332; Cell Signaling) was used. Following incubation, membranes were washed 3 times with TBST before incubation with an appropriate peroxidase-conjugated secondary antibody in TBST (1:10,000 dilution). Antibody binding was detected using the enhanced chemiluminescence horseradish peroxidase (HRP) substrate detection kit (32106; Pierce).

Microarray and bioinformatic analysis. RNA from gastrocnemius was isolated with an miRNeasy minikit (Qiagen), and microarray was performed using the GeneChip Gene 1.0 ST Array System (Affymetrix). In addition, gene ontology (GO) analysis was performed with the online tool FatiGO (http://babelomics.bioinfo.cipf.es/index.html) (2). Finally, microarray data were analyzed using Motif Activity Response Analysis (MARA; http://www.mara.unibas.ch/cgi/mara) (47).

Statistical analysis. Values are expressed as means ± standard errors of the means (SEM). Statistical significance was determined with unpaired two-tailed *t* tests or one-way analysis of variance (ANOVA) with Tukey's *post hoc* test. A *P* value <0.05 was considered significant.

Microarray data accession number. Microarray data can be found at the Gene Expression Omnibus (GEO) under accession number GSE40439.

RESULTS

Muscle NCoR1 deletion enhances VO_2 during maximal exercise and decreases muscle contractility. The role of NCoR1 in skeletal muscle was studied using NCoR1 MKO mice. A full description of this animal model has been recently published elsewhere (51). First, we assessed maximal oxidative capacity and treadmill running performance. Both genotypes exhibited the same exercise performance during maximal and endurance exercise tests, as reflected by equal speed, distance, time, work, and power (see Table S1 in the supplemental material). However, NCoR1 MKO mice reached a significantly higher (7.2%) VO_{2peak} during the maximal exercise test (CON mice, 125 ± 1.2 ml/kg of body weight/min; MKO mice, 134 ± 1.5 ml/kg/min; $P < 0.001$). In fact, NCoR1 MKO animals showed a higher VO_2 at 50, 80, 90, and 100% of maximal speed (Fig. 1A). Blood lactate measurement before and after maximal exercise, or during fasting, was not different in NCoR1 MKO mice (see Fig. S2A and B in the supplemental material). Consistent with this, respiratory exchange ratio (RER) measurement during the maximal exercise test did not show differences between CON and NCoR1 MKO mice (see Fig. S2C), indicating that energy substrate utilization was not altered by NCoR1 deletion. We then performed NADH and PAS staining of skeletal muscle to determine the proportion of oxidative fibers and glycogen content, respectively. NCoR1 MKO animals showed a 10% increase in oxidative fibers compared to CON mice (Fig. 1B), but glycogen content was not significantly affected (Fig. 1C). Furthermore, we have found higher levels of AMP-activated protein kinase (AMPK) phosphorylation in skeletal muscle of NCoR1 MKO mice (Fig. 1D). Altogether, these data show that NCoR1 deletion in striated muscle results in an enhanced oxidative metabolism, particularly during high-intensity exercise.

An important aspect of muscle function besides endurance is the ability to generate force. To assess muscle contractility *in vivo*, we measured maximal grip strength of the fore and hind limbs. We observed a trend for lower peak isometric force generated by the fore limbs ($P = 0.078$) and a significant reduction by 11% in the hind limbs of NCoR1 MKO animals (Fig. 1E). We then quantified muscle fatigue resistance during predominantly isometric muscle contraction by measuring the maximal time that mice could remain on an inverted grid. Consistent with the lower maximal isometric force, isometric fatigue resistance was also decreased in NCoR1 MKO animals (Fig. 1F). To exclude systemic and neural factors that could affect force generation *in vivo*, we also determined muscle contractility in isolated muscles. Absolute ($P < 0.05$) and specific ($P = 0.056$) muscle contractility in response to a single twitch was decreased in the glycolytic muscle extensor digitorum longus (EDL) but not in the oxidative muscle soleus (Fig. 2A and B). Conversely, when EDL and soleus were subjected to maximal tetanic stimulation, no differences between genotypes were observed (Fig. 2C and D). Moreover, the contractile kinetics of EDL and soleus in response to a single twitch or a maximal tetanic stimulation were not significantly different be-

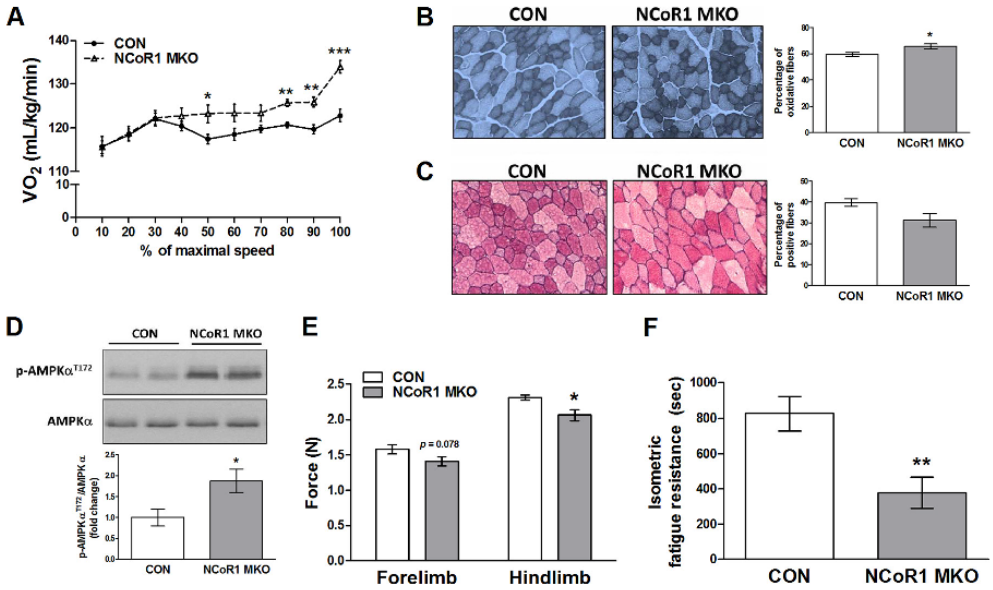


FIG 1 Exercise performance and *in vivo* contractile properties of NCoR1 MKO mouse skeletal muscle. (A) VO_2 during the maximal exercise test ($n = 8$ CON and $n = 7$ NCoR1 MKO mice). (B and C) Representative pictures and quantification of NADH (B) and PAS (C) staining of tibialis anterior ($n = 3$ to 4 CON and $n = 3$ to 4 NCoR1 MKO mice). (D) Representative blots and quantification of gastrocnemius AMPK phosphorylation levels ($n = 7$ CON and $n = 4$ NCoR1 MKO mice). (E and F) *In vivo* assessment of maximal grip strength and isometric muscle fatigue resistance ($n = 13$ CON and $n = 12$ NCoR1 MKO mice). Values represent means ± SEM. $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (**) for CON versus NCoR1 MKO mice.

tween CON and NCoR1 MKO mice (see Table S1 in the supplemental material).

Finally, muscle fatigue resistance was determined *ex vivo* by a long- and short-interval protocol. As expected, soleus exhibited a lower decrease in force than EDL in the long-interval protocol, but no differences between genotypes were found (Fig. 2E). Similarly, soleus had higher fatigue resistance than EDL during the short-interval protocol. However, in this protocol, EDL from NCoR1 MKO animals exhibited a lower decrease in force generation, while soleus of NCoR1 MKO and CON mice were identical (Fig. 2F). In fact, in the short-interval protocol, EDL from NCoR1 MKO mice generated ~43% more force from tetanic stimulation no. 35 to 100 (Fig. 2F, inset), indicating a higher muscle fatigue resistance. Therefore, it seems that glycolytic muscles are more susceptible to the effect of NCoR1 deletion on muscle contractility, resulting in a decreased maximal isometric force and increased fatigue resistance during maximal stimulation.

NCoR1 and PGC-1α target a common subset of genes involved in oxidative metabolism. The phenotype exhibited by NCoR1 MKO animals implies a direct link between NCoR1 and oxidative metabolism. At the transcriptional level, mitochondrial function is mainly regulated by the PPARγ coactivator 1 (PGC-1) family of coactivators, which are able to activate different transcription factors, such as ERRα and NRF-1 (29). Interestingly, NCoR1 MKO mice mirror some aspects of the phenotype exhibited by PGC-1α mTg mice, such as the enhanced oxidative metabolism and decreased maximal force (11, 30, 45). To explore the idea of an NCoR1-PGC-1α cross talk and to further characterize the oxidative phenotype of NCoR1 MKO mice, microarray analyses of gene expression patterns in NCoR1 MKO and PGC-1α mTg skeletal muscles were compared. GO enrichment analysis revealed overrepresentation of transcripts related to metabolic pathways, such as oxidative phosphorylation and the citrate cycle (tricarboxylic acid [TCA] cycle), in both NCoR1 MKO and PGC-1α mTg animals (Fig. 3A and B; also see Fig. S4A to F in the supplemental material). Interestingly, when both microarray data sets were compared, we observed that 188 of the genes affected by NCoR1 deletion were also found in the PGC-1α mTg data set, suggesting that ~50% of the genes regulated by NCoR1 are also targets of PGC-1α (Fig. 3C). When only the genes found with the GO terms oxidative phosphorylation and citrate cycle (TCA cycle) were compared, we found that all of the genes present in the NCoR1 MKO data set were also targets of PGC-1α (Fig. 3C). Interestingly, we observed that in both microarrays all of these common genes were actually upregulated (Fig. 3D and E). Several of these transcripts were validated by real-time PCR analysis of GAS, EDL, and soleus (see Fig. S3B and C in the supplemental material), confirming the results obtained with the microarrays. However, no fiber type-specific (oxidative versus glycolytic) differences in terms of regulation of gene expression were observed in NCoR1 MKO mice (see Fig. S3C). Furthermore, skeletal muscle from NCoR1 MKO and PGC-1α mTg mice had higher transcript levels of mitochondrial genes involved in oxidative metabolism, such as cytochrome *c* oxidase subunit I (*COX1*), ATP synthase F0 subunit 6 (*ATP6*), and NADH dehydrogenase subunit 1 (*ND1*) (Fig. 3F; also see Fig. S3B). However, no differences in transcription factor A, mitochondrial (*TFAM*) mRNA levels, and mitochondrial DNA content were found as a consequence of NCoR1 deletion (Fig. 3G; also see Fig. 6A). These results further demonstrate the role of NCoR1 in the transcriptional control of different mitochondrial

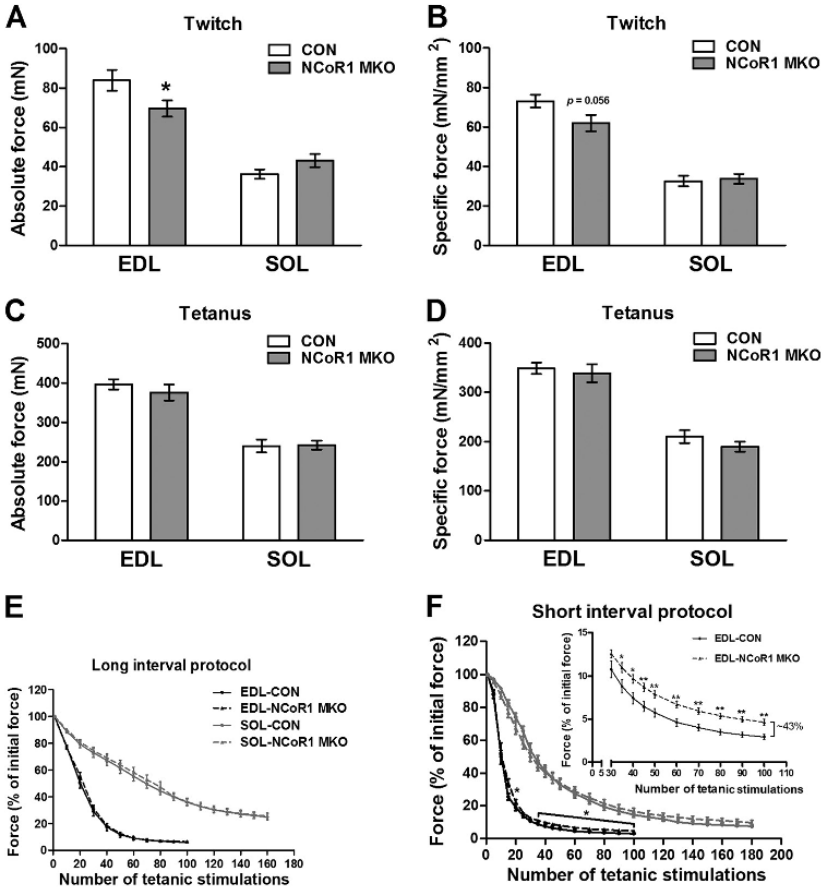
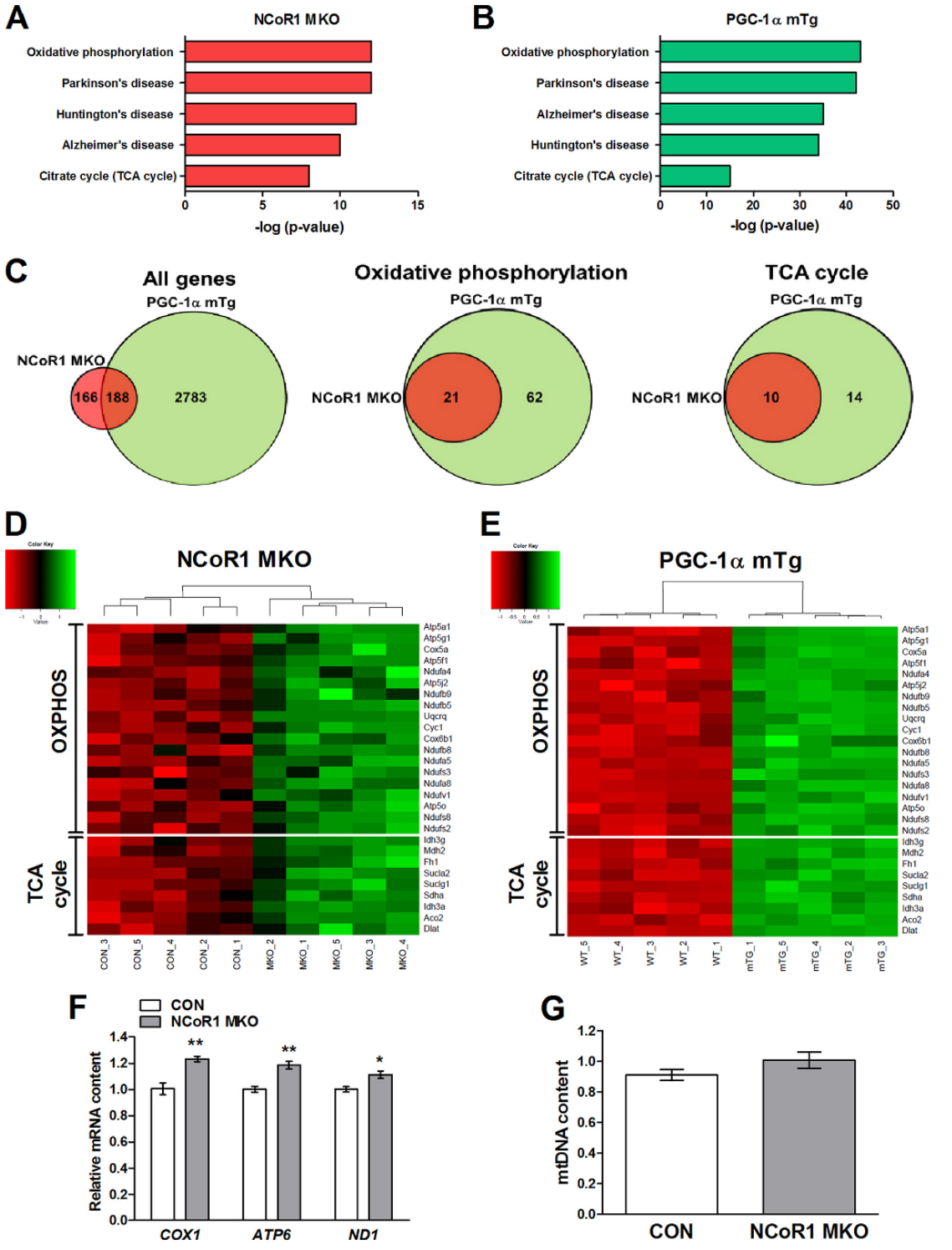


FIG 2 *Ex vivo* contractile properties of NCoR1 MKO mouse skeletal muscle. (A to D) *Ex vivo* assessment of extensor digitorum longus (EDL) and soleus (SOL) contractility in response to a single twitch or a maximal tetanic stimulation ($n = 10$ CON and $n = 14$ NCoR1 MKO muscles per group). (E and F) *Ex vivo* assessment of EDL and SOL fatigue resistance in response to repeated tetanic stimulation through a long- and short-interval protocol. The inset in panel F shows the section with significant differences ($n = 10$ CON and $n = 14$ MKO muscles per group). Values represent means ± SEM. $P < 0.05$ (*) and $P < 0.01$ (**) for CON versus MKO mice.

pathways and strongly suggest a role of PGC-1α in the control of the oxidative phenotype of the NCoR1 MKO animals.

Interestingly, the mRNA level of PGC-1α was not different between NCoR1 MKO and CON mice (Fig. 4A; also see Fig. S5A in the supplemental material). In contrast, we found an increase in PGC-1β and a slight decrease in PGC-1-related coactivator (PRC) transcript levels in gastrocnemius and soleus from NCoR1 MKO mice (Fig. 4A; also see Fig. S5A). To dissect the contribution of the PGC-1 family members to the effects induced by NCoR1 knockdown, we studied C_2C_{12} myotubes in culture that were virally transfected with two different shRNAs specifically targeting NCoR1 (Fig. 4B; also see Fig. S5B and C and S6A and B in the supplemental material). We then selectively induced the knockdown of PGC-1α or PGC-1β with specific shRNA constructs (Fig. 4B; also see Fig. S5B). Similar to the NCoR1 MKO animals, NCoR1 knockdown in C_2C_{12} myotubes also induced an upregulation of succinate dehydrogenase complex, subunit A, flavoprotein (*SDHa*), NADH dehydrogenase 1α (ubiquinone) subcomplex 5 (*NDUFA5*), NADH dehydrogenase 1β (ubiquinone) subcomplex 5 (*NDUFB5*), and fumarate hydratase 1 (*FH1*) (Fig. 4C), in addition to the mitochondrial genes *COX1* and *ATP6* (Fig. 4C). Surprisingly, NCoR1 knockdown in C_2C_{12} myotubes induced a big increase in PGC-1α mRNA (Fig. 4B). Importantly, however, although the simultaneous knockdown of PGC-1α partially prevented its upregulation, shRNA-mediated reduction of PGC-1α was sufficient to strongly inhibit the upregulation of *SDHa*, *NDUFA5*, *NDUFB5*, *FH1*, *COX1*, and *ATP6* mRNA induced by NCoR1 knockdown (Fig. 4C). In the cultured muscle cells, NCoR1 knockdown did not alter PGC-1β gene expression (see Fig. S5B), while the opposite was observed in the NCoR1 MKO animals (Fig. 4A; also see Fig. S5A). In stark contrast to the knockdown of PGC-1α, the upregulation of the nuclear and mitochondrial genes induced by NCoR1 knockdown was not significantly affected by the simultaneous knockdown of PGC-1β in the cultured myotubes (Fig. 4D). Consistent with the microarray analysis, these results indicate


A
NCoR1 MKO
Oxidative phosphorylation
Parkinson's disease
Huntington's disease
Alzheimer's disease
Citrate cycle (TCA cycle)
0
5
10
15
-log (p-value)
B
PGC-1α mTg
Oxidative phosphorylation
Parkinson's disease
Alzheimer's disease
Huntington's disease
Citrate cycle (TCA cycle)
0
10
20
30
40
50
-log (p-value)
C
All genes
PGC-1α mTg
NCoR1 MKO
166
188
2783
Oxidative phosphorylation
PGC-1α mTg
NCoR1 MKO
21
62
TCA cycle
PGC-1α mTg
NCoR1 MKO
10
14
D
NCoR1 MKO
OXPHOS
TCA cycle
Atp5a1
Atp5g1
Cox5a
Atp5f1
Ndufa4
Atp5j2
Ndufb9
Ndufb5
Uqcrq
Cyc1
Cox6b1
Ndufb8
Ndufa5
Ndufs3
Ndufa8
Ndufv1
Atp5o
Ndufs8
Ndufs2
Idh3g
Mdh2
Fh1
Sucla2
Suclg1
Sdha
Idh3a
Aco2
Dlat
CON_3
CON_5
CON_4
CON_2
CON_1
MKO_2
MKO_1
MKO_5
MKO_3
MKO_4
E
PGC-1α mTg
OXPHOS
TCA cycle
WT_5
WT_4
WT_3
WT_2
WT_1
mTG_1
mTG_5
mTG_4
mTG_2
mTG_3
F
CON
NCoR1 MKO
Relative mRNA content
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
**
**
*
COX1
ATP6
ND1
G
mtDNA content
0.0
0.2
0.4
0.6
0.8
1.0
1.2
CON
NCoR1 MKO

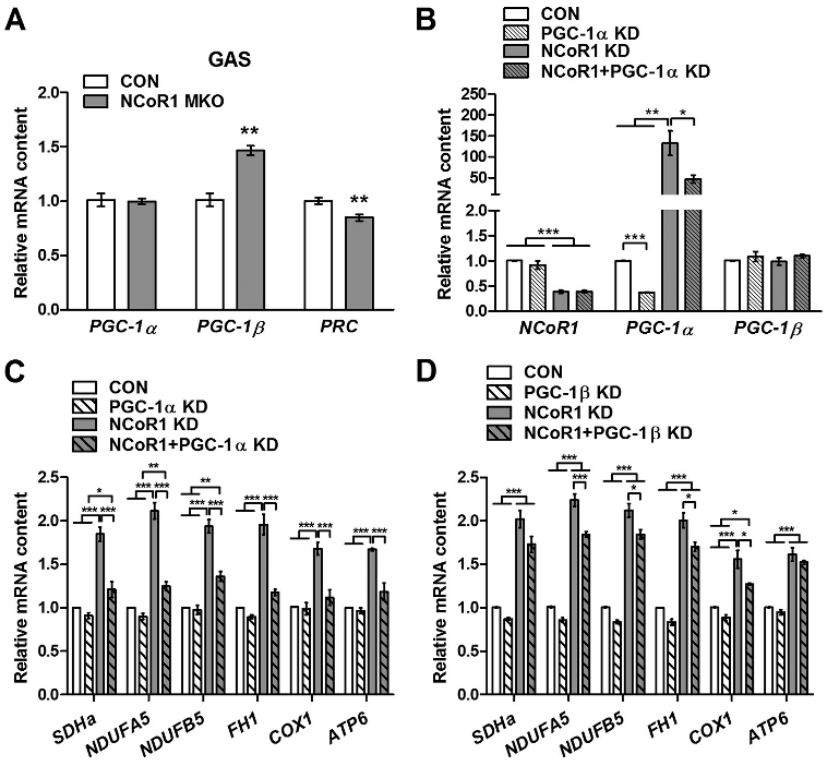


FIG 4 Role of PGC-1α/β in NCoR1 regulation of oxidative metabolism. (A) Gastrocnemius (GAS) mRNA levels of the PGC-1 family of coactivators ($n = 7$ CON and $n = 5$ NCoR1 MKO mice). (B to D) Adenoviral knockdown (KD) of LacZ (CON) or NCoR1 alone or in combination with PGC-1α or PGC-1β KD for 48 h in C_2C_{12} myotubes ($n = 3$ independent experiments each performed in triplicate). Values represent means ± SEM. $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***) for CON versus MKO mice or as indicated.

that NCoR1 deletion enhances PGC-1α action on oxidative metabolism, probably due to a reduced competition between NCoR1 and PGC-1α for the regulation of common target genes.

NCoR1 and PGC-1α regulate oxidative metabolism through opposite modulation of ERRα. Our data suggest that NCoR1 and PGC-1α target the same subset of transcription factors involved in the regulation of oxidative metabolism with opposite effects on transcriptional activation. In order to elucidate potential common transcription factors of NCoR1 and PGC-1α in skeletal muscle, we performed motif activity response analysis (MARA) (47) of our microarray data to predict the core set of transcription factors which significantly change their activity in response to NCoR1 deletion or PGC-1α overexpression. Consistent with the oxidative phenotype of these two mouse models, retinoid X receptor (RXR) and ERRα (also known as *Esrra*) were among the top 4 most significant motifs (z value, ≥1.5) (Fig. 5A and B; also see Table S2 in the supplemental material) and exhibited some of the highest levels of activity in response to NCoR1 deletion and PGC-1α overexpression (Fig. 5C to F). Given that RXRs are the main partners of PPARs (13), this suggests possible PPARβ/δ activation in MKO animals. In fact, PPARβ/δ or ERRα in complex with PGC-1α and PGC-1β are known to be key players in the regulation of muscle metabolism (13, 16, 29). In contrast, the role of the transcription factors with significantly decreased activity (z value, ≤−1.5) in NCoR1 MKO and PGC-1α mTg mice (Fig. 5A and B; also see Table S2) is not well known. Similar to the microarray analysis, the comparison of both MARA performances showed that 44% of the transcription factors found in the NCoR1 MKO data set were also predicted in PGC-1α mTg MARA; there is actually a larger overlap among motifs with increased activity (5 out of 9) than among ones with decreased activity (2 out of 7) (Fig. 5G). In summary, MARA-

FIG 3 Similarities between NCoR1 MKO and PGC-1α mTg mice on oxidative metabolism. (A and B) Top 5 KEGG pathways from GO analysis of the up- and downregulated genes from the NCoR1 MKO and PGC-1α mTg microarray data sets ($n = 5$ per group). (C) Venn diagrams showing the overlap between NCoR1 MKO and PGC-1α mTg microarray data sets. (D and E) Heat maps generated using probe set intensities of the overlapping transcripts between NCoR1 MKO and PGC-1α mTg related to the GO terms oxidative phosphorylation (OXHPOS) and citrate cycle (TCA cycle). (F) Real-Time PCR analysis of mRNA levels of mitochondrial genes of gastrocnemius from CON ($n = 7$) and NCoR1 MKO ($n = 5$) animals. (G) Measurement of mitochondrial DNA (mtDNA) content of gastrocnemius from CON ($n = 7$) and NCoR1 MKO ($n = 5$) animals. Values represent means ± SEM. $P < 0.05$ (*) and $P < 0.01$ (**) for CON versus NCoR1 MKO mice.

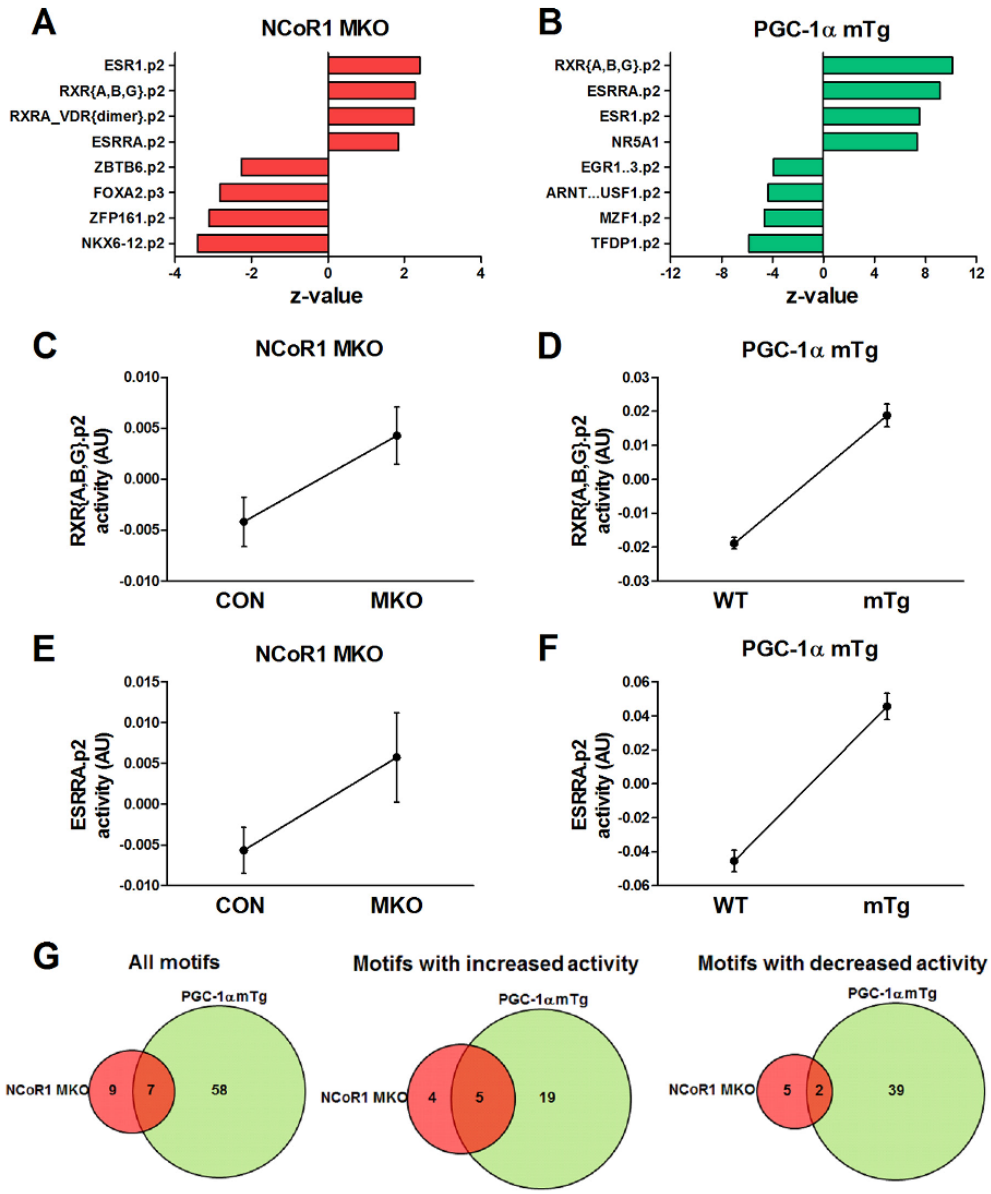


FIG 5 Identification of common transcription factor binding partners of NCoR1 and PGC-1α. (A and B) The top 4 transcription factor motifs exhibiting increased and decreased activity in MARA of the microarray data performed on gastrocnemius from NCoR1 MKO and PGC-1α mTg mice ($n = 5$ per group). (C to F) Changes in the activity of RXRs and ERRα (ESRRA) in NCoR1 MKO and PGC-1α mTg skeletal muscle predicted by MARA analysis of microarray data ($n = 5$ per group). (G) Venn diagrams showing the overlap between NCoR1 MKO and PGC-1α mTg MARA analysis.

based biocomputational prediction strongly suggests ERRα and possibly the RXR heterodimerization partner PPARβ/δ as common targets of NCoR1 and PGC-1α.

The mRNA levels of PPARβ/δ, ERRα, and several transcription factors revealed by MARA were not increased in NCoR1 MKO mice (Fig. 6A), indicating that increased activity is the consequence of lower repression rather than of higher expression. In order to explore the potential repressive effect of NCoR1 on ERRα and PPARβ/δ transcriptional activity, we transfected COS-7 cells with a reporter plasmid containing PPAR response elements (PPRE-luc) or ERR response elements (ERRE-luc) together with expression plasmids for PPARβ/δ and ERRα, respectively. We

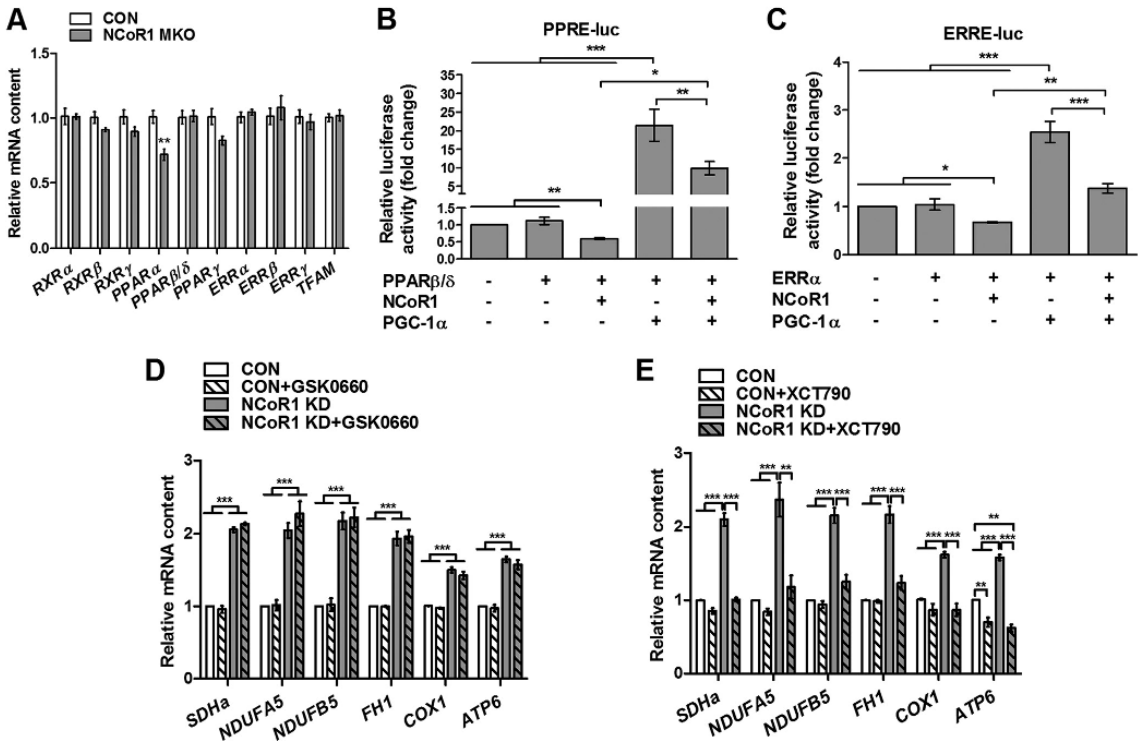


FIG 6 Role of PPARβ/δ and ERRα in NCoR1 regulation of oxidative metabolism. (A) Gastrocnemius mRNA levels of different transcription factors ($n = 7$ CON and $n = 5$ NCoR1 MKO mice). RXR, retinoid X receptor; PPAR, peroxisome proliferator activated receptor; ERR, estrogen-related receptor; TFAM, transcription factor A, mitochondrial. (B and C) Luciferase activity of PPRE-luc and ERRE-luc reporter plasmids in COS-7 cells cotransfected with PPARβ/δ, ERRα, NCoR1, and PGC-1α as indicated in each figure ($n = 5$ independent experiments each performed in triplicate). (D and E) Adenoviral knockdown (KD) of LacZ (CON) or NCoR1 alone or in combination with 10 μM XCT790 or 1 μM GSK0660 for 48 h in C_2C_{12} myotubes ($n = 3$ independent experiments each performed in triplicate). Values represent means ± SEM. $P < 0.05$ (*), $P < 0.01$ (**), and $P < 0.001$ (***) for CON versus NCoR1 MKO mice or as indicated.

next measured relative luciferase activity in the absence and presence of NCoR1 and PGC-1α. As predicted, NCoR1 decreased PPRE-luc and ERRE-luc luciferase activity by 43 and 36%, respectively (Fig. 6B and C). Inversely, PGC-1α induced a significant increase in PPRE-luc and ERRE-luc luciferase activity, 2,037 and 161%, respectively (Fig. 6B and C). Importantly, we have found that the activation of both PPRE-luc and ERRE-luc by PGC-1α was significantly decreased by NCoR1 (Fig. 6B and C). These data demonstrate that NCoR1 represses the transcriptional activity of both PPARβ/δ and ERRα, while PGC-1α competes with NCoR1 to exert a positive effect.

Finally, to study the relative contribution of PPARβ/δ and ERRα to the regulation of the oxidative phenotype exhibited in response to NCoR1 deletion in muscle, we used the PPARβ/δ selective antagonist GSK0660 (44) and the ERRα inverse agonist XCT790 (33, 49). As expected, GSK0660 induced a strong decrease in the mRNA level of the PPARβ/δ target gene uncoupling protein 3 (*UCP3*) (see Fig. S6A in the supplemental material), demonstrating the efficiency of the antagonist and the presence of functional PPARβ/δ in C_2C_{12} myotubes. Neither NCoR1 knockdown nor GSK0660 treatment changed *PPARβ/δ* mRNA levels (see Fig. S6A). Importantly, however, the induction of *SDHa*, *NDUFA5*, *NDUFB5*, *FH1*, *COX1*, and *ATP6* mRNA were not inhibited by GSK0660 in cells with a knockdown of NCoR1 (Fig. 6D). Consistent with this, the mRNA levels of different PPARβ/δ target genes, like the carnitine palmitoyltransferase 1b gene (*CPT1b*), the lipoprotein lipase gene (*LPL*), and *UCP3*, were not significantly increased in NCoR1 MKO mice (see Fig. S3C in the supplemental material). In contrast, ERRα inhibition with XCT790 completely blocked the effects of NCoR1 knockdown on nuclear and mitochondrial genes (Fig. 6E). However, unlike the NCoR1 MKO animals, NCoR1 knockdown induced an upregulation of *ERRα* in C_2C_{12} myotubes (see Fig. S6B). This upregulation of *ERRα* mRNA was only partially prevented by XCT790 (see Fig. S6B), indicating that ERRα activation, rather than its upregulation, is responsible for the effects of NCoR1 knockdown of oxidative metabolism in C_2C_{12} myotubes. Interestingly, similar to NCoR1 and PGC-1α microarray analysis, GO enrichment analysis of previously published ERRα chromatin immunoprecipitation (ChIP)-on-chip data from mouse liver (12) also revealed overrepresentation of transcripts related to oxidative phosphorylation (see Fig. S4G in the supplemental material). Furthermore, we found a high overlap between the NCoR1 MKO microarray and the ERRα ChIP-on-chip data sets when comparing the genes found with the GO term oxidative phosphorylation (see Fig. S4H). Therefore, these data suggest that ERRα and PGC-1α are essential for the effects of NCoR1 deletion on oxidative metabolism in

skeletal muscle, while PPARβ/δ and PGC-1β seem to play minor roles in this experimental context.

DISCUSSION

In stark contrast to obese and type 2 diabetic subjects, endurance athletes exhibit an increased metabolic fitness and consequently lower risk for metabolic disorders (9, 35). Importantly, most of the adaptations to muscle use (e.g., endurance training) and disuse (e.g., physical inactivity) are under the coordinated control of different transcription factors and coregulators (7, 19). NCoR1 and its homolog NCoR2 have been recently suggested as important modulators of energy metabolism in several tissues, including skeletal muscle (28, 40, 46, 51). Consistent with this, we observed that NCoR1 MKO animals exhibit an increased VO_2 during high-intensity exercise. Surprisingly, while Yamamoto et al. (51) have reported improved exercise performance in NCoR1 MKO mice, we did not observe significant differences in distance, time, or work in the exercise trial in our experimental context. Considering the mild effect of NCoR1 deletion on muscle oxidative metabolism, it is possible that small differences in the exercise test protocol significantly affect exercise performance. In addition, as previously shown (51), we observed a higher proportion of oxidative fibers in NCoR1 MKO skeletal muscle. Interestingly, our data also indicate a higher activation of AMPK in NCoR1 MKO mice, which has been associated with enhanced oxidative metabolism (20, 21). Importantly, we have now demonstrated that NCoR1 controls skeletal muscle oxidative metabolism primarily through the regulation of ERRα and PGC-1α.

NCoR1 deletion in striated muscle led to an increase in the mRNA content of a broad range of mitochondrial enzymes, thus supporting its potential role as a negative regulator of oxidative metabolism. In agreement with this idea, NCoR1 MKO mice recapitulate many aspects of the phenotype exhibited by PGC-1α mTg mice, such as the increased expression of mitochondrial enzymes, higher levels of oxidative fibers, and enhanced VO_{2peak} during maximal exercise (11, 30, 51). Moreover, as found in NCoR1 MKO animals, PGC-1α overexpression in skeletal muscle also results in a reduced maximal force and increased muscle fatigue resistance (30, 45). Curiously, only EDL exhibited a higher muscle fatigue resistance during the short-interval protocol in NCoR1 MKO mice. Since EDL is a highly glycolytic muscle, it might be more susceptible to benefitting from the mild improvement in oxidative metabolism as a consequence of NCoR1 deletion compared to the predominantly oxidative soleus. Analysis of NCoR1 MKO- and PGC-1α mTg-based microarrays also showed a high level of similarity between these mouse models, since all of the upregulated genes related to oxidative metabolism found in NCoR1 MKO mice were also increased by PGC-1α overexpression. Hence, our findings indicate that NCoR1 deletion facilitates PGC-1α action on its target genes. In fact, PGC-1α mRNA was not increased in NCoR1 MKO mice, suggesting that the higher transcription of its target genes is not associated with its upregulation. In contrast, NCoR1 knockdown in cultured cells induced a 133-fold increase in PGC-1α mRNA. Hence, it is possible that the bigger effects of NCoR1 deletion in C_2C_{12} myotubes compared to those on the mice are a consequence of both upregulation of PGC-1α and lower competition between these coregulators. Importantly, however, we found that knockdown of PGC-1α, but not PGC-1β, in C_2C_{12} myotubes completely blocked the effects of NCoR1 knockdown on oxidative metabolism, underlining the essential role of PGC-1α in this process. Similarly, prediction of the relative contribution of transcription factors to the gene expression pattern exhibited by NCoR1 MKO and PGC-1α mTg animals revealed a significant overlap between these mouse models. Interestingly, RXRs (heterodimerization partners for the PPARs) and ERRα stood out as top candidates, showing the strongest link to the control of energy metabolism (13, 16). In fact, Yamamoto et al. (51) have recently showed that in C_2C_{12} myotubes, NCoR1 is recruited to the PPRE and nuclear receptor half-sites (ERR binding site) of *UCP3* and *PDK4* promoter, respectively, though whether the effects of NCoR1 deletion on gene expression actually depend on PPARβ/δ or ERRα activation was not studied. Here, through reporter gene assays, we have demonstrated that NCoR1 is a direct corepressor of both PPARβ/δ and ERRα. Importantly, given that NCoR1 and PGC-1α compete for the transcriptional regulation of PPARβ/δ and ERRα, our data indicate that these transcription factors are common targets of both coregulators. Interestingly, we observed a striking effect of the ERRα inverse agonist XCT790, since it completely inhibited the increase in the mRNA level of several mitochondrial enzymes triggered by NCoR1 knockdown in cultured myotubes. These data are consistent with MARA results and suggest that ERRα is the main target of NCoR1 in the control of muscle oxidative metabolism (Fig. 7).

Unexpectedly, we found that the PPARβ/δ selective antagonist GSK0660 did not affect the effects of NCoR1 deletion on oxidative phosphorylation and gene regulation. PPARβ/δ plays an impor-

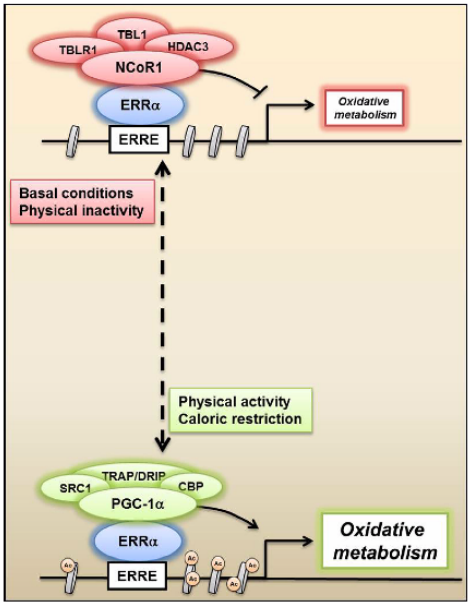


FIG 7 Antagonistic regulation of oxidative metabolism by NCoR1 and PGC-1α. Proposed mechanism by which NCoR1 and PGC-1α complexes compete for the transcriptional regulation of ERRα and, as a consequence, oxidative metabolism. Under basal and possibly pathological conditions, NCoR1 represses ERRα and increases histone deacetylation, thereby decreasing metabolic gene transcription. In contrast, PGC-1α is activated by exercise and caloric restriction in skeletal muscle, coactivates ERRα, and consequently enhances mitochondrial function. ERRE, ERR response elements; Ac, histone acetylation.

tant role in the transcriptional control of lipid metabolism (13). Intriguingly, our microarray analysis of NCoR1 MKO skeletal muscle shows a preferential enhancement of oxidative phosphorylation and the TCA cycle compared to the modulation of fatty acid β-oxidation. Thus, it is conceivable that in our experimental context of animals fed with normal chow, PPARβ/δ played less of a role than it did in studies using high-fat diets (51) or under conditions like obesity and physical inactivity. NCoR1 might control fatty acid metabolism through the transcriptional regulation of PPARβ/δ in an environment of elevated intramyocellular lipids that could act as PPARβ/δ ligands, but this hypothesis needs to be substantiated in future studies. Overall, it seems that a fully functional ERRα-PGC-1α complex is a prerequisite for the improved oxidative metabolism induced by muscle-specific NCoR1 deletion, at least in chow-fed mice.

Interestingly, NCoR1 MKO mice have a milder oxidative phenotype than PGC-1α transgenic mice. The weaker phenotype is also reflected in the quantitative differences in gene expression. Thus, while the muscle knockout of NCoR1 resulted in the upregulation of 21 genes involved in oxidative phosphorylation, 83 genes were increased in response to muscle PGC-1α overexpression. Consistent with this, PGC-1α mTg mice exhibit a higher increase in VO_{2peak} (24%) (11) than NCoR1 MKO animals (7.2%). However, it is important to note that NCoR1 is a basal corepressor, thus it represses transcription factors under basal conditions, and it is exchanged by coactivators upon a positive stimulus (17, 37). Therefore, our data suggest that under basal conditions, NCoR1 deletion would not result in a full activation of PPARβ/δ and ERRα, which subsequently could be amplified by a positive stimulus, such as exercise, that is well known to increase PGC-1α expression (4) and thereby unleashes the full activity of these transcription factors.

In conclusion, our data indicate competition between NCoR1 and PGC-1α in the regulation of PPARβ/δ and ERRα transcriptional activity in skeletal muscle, of which the regulation of ERRα represents the predominant regulatory mechanism of oxidative metabolism during basal conditions (Fig. 7). The elucidation of different pharmacological or nonpharmacological (e.g., exercise training) strategies to modulate NCoR1 activity and thus facilitate PGC-1α action represents an attractive strategy for the treatment of metabolic diseases.

ACKNOWLEDGMENTS

We thank Christopher K. Glass (University of California, San Diego), Junichi Sadoshima (University of Medicine and Dentistry of New Jersey), and Vincent Giguère (McGill University, Montreal, Canada) for their generous contribution of pFlag-NCoR, pERRE-luc, and pERRα plasmids, respectively. We also thank Philippe Demougin (Biozentrum, University of Basel) for his help with microarray analysis and Hiroyasu Yamamoto for providing the LacZ and NCoR1 adenoviruses (Ecole Polytechnique Fédérale de Lausanne).

This project was funded by the Swiss National Science Foundation (SNF 310030_132900), the Muscular Dystrophy Association USA, the SwissLife Jubiläumsstiftung für Volksgesundheit und medizinische Forschung, the Swiss Society for Research on Muscle Diseases (SSEM), the Swiss Diabetes Association, the Roche Research Foundation, the United Mitochondrial Disease Foundation, the Association Française contre les Myopathies, the Ecole Polytechnique Féderale de Lausanne, the European Research Council, and the University of Basel.

REFERENCES

1. **Alenghat T, et al.** 2008. Nuclear receptor corepressor and histone deacetylase 3 govern circadian metabolic physiology. Nature **456**:997–1000.
2. **Al-Shahrour F, Diaz-Uriarte R, Dopazo J.** 2004. FatiGO: a web tool for finding significant associations of gene ontology terms with groups of genes. Bioinformatics **20**:578–580.
3. **Baar K.** 2010. Epigenetic control of skeletal muscle fibre type. Acta Physiol. **199**:477–487.
4. **Baar K, et al.** 2002. Adaptations of skeletal muscle to exercise: rapid increase in the transcriptional coactivator PGC-1. FASEB J. **16**:1879–1886.
5. **Bailey P, et al.** 1999. The nuclear receptor corepressor N-CoR regulates differentiation: N-CoR directly interacts with MyoD. Mol. Endocrinol. **13**:1155–1168.
6. **Barres R, et al.** 2012. Acute exercise remodels promoter methylation in human skeletal muscle. Cell Metab. **15**:405–411.
7. **Bassel-Duby R, Olson EN.** 2006. Signaling pathways in skeletal muscle remodeling. Annu. Rev. Biochem. **75**:19–37.
8. **Booth FW, Lees SJ.** 2007. Fundamental questions about genes, inactivity, and chronic diseases. Physiol. Genomics **28**:146–157.
9. **Booth FW, Roberts CK.** 2008. Linking performance and chronic disease risk: indices of physical performance are surrogates for health. Br. J. Sports Med. **42**:950–952.
10. **Busson M, et al.** 2005. Coactivation of nuclear receptors and myogenic factors induces the major BTG1 influence on muscle differentiation. Oncogene **24**:1698–1710.
11. **Calvo JA, et al.** 2008. Muscle-specific expression of PPARgamma coactivator-1alpha improves exercise performance and increases peak oxygen uptake. J. Appl. Physiol. **104**:1304–1312.
12. **Charest-Marcotte A, et al.** 2010. The homeobox protein Prox1 is a negative modulator of ERRα/PGC-1α bioenergetic functions. Genes Dev. **24**:537–542.
13. **Ehrenborg E, Krook A.** 2009. Regulation of skeletal muscle physiology and metabolism by peroxisome proliferator-activated receptor delta. Pharmacol. Rev. **61**:373–393.
14. **Fischle W, et al.** 2002. Enzymatic activity associated with class II HDACs is dependent on a multiprotein complex containing HDAC3 and SMRT/N-CoR. Mol. Cell **9**:45–57.
15. **Furuya F, et al.** 2007. Nuclear receptor corepressor is a novel regulator of phosphatidylinositol 3-kinase signaling. Mol. Cell. Biol. **27**:6116–6126.
16. **Giguere V.** 2008. Transcriptional control of energy homeostasis by the estrogen-related receptors. Endocr. Rev. **29**:677–696.
17. **Glass CK, Rosenfeld MG.** 2000. The coregulator exchange in transcriptional functions of nuclear receptors. Genes Dev. **14**:121–141.
18. **Gonzalez E, Delbono O.** 2001. Age-dependent fatigue in single intact fast and slow fibers from mouse EDL and soleus skeletal muscles. Mech. Ageing Dev. **122**:1019–1032.
19. **Handschin C, Spiegelman BM.** 2008. The role of exercise and PGC1alpha in inflammation and chronic disease. Nature **454**:463–469.
20. **Hardie DG, Ross FA, Hawley SA.** 2012. AMPK: a nutrient and energy sensor that maintains energy homeostasis. Nat. Rev. Mol. Cell Biol. **13**:251–262.
21. **Hawley SA, et al.** 2012. The ancient drug salicylate directly activates AMP-activated protein kinase. Science **336**:918–922.
22. **Horlein AJ, et al.** 1995. Ligand-independent repression by the thyroid hormone receptor mediated by a nuclear receptor co-repressor. Nature **377**:397–404.
23. **Ishizuka T, Lazar MA.** 2005. The nuclear receptor corepressor deacetylase activating domain is essential for repression by thyroid hormone receptor. Mol. Endocrinol. **19**:1443–1451.
24. **Jepsen K, et al.** 2000. Combinatorial roles of the nuclear receptor corepressor in transcription and development. Cell **102**:753–763.
25. **Kelley DE, He J, Menshikova EV, Ritov VB.** 2002. Dysfunction of mitochondria in human skeletal muscle in type 2 diabetes. Diabetes **51**:2944–2950.
26. **Kitamura T, et al.** 2007. A Foxo/Notch pathway controls myogenic differentiation and fiber type specification. J. Clin. Investig. **117**:2477–2485.
27. **Kouzarides T.** 2007. Chromatin modifications and their function. Cell **128**:693–705.
28. **Li P, et al.** 2011. Adipocyte NCoR knockout decreases PPARgamma

phosphorylation and enhances PPARgamma activity and insulin sensitivity. Cell **147**:815–826.
29. **Lin J, Handschin C, Spiegelman BM.** 2005. Metabolic control through the PGC-1 family of transcription coactivators. Cell Metab. **1**:361–370.
30. **Lin J, et al.** 2002. Transcriptional co-activator PGC-1 alpha drives the formation of slow-twitch muscle fibres. Nature **418**:797–801.
31. **McGee SL, Hargreaves M.** 2011. Histone modifications and exercise adaptations. J. Appl. Physiol. **110**:258–263.
32. **McKenna NJ, O'Malley BW.** 2010. SnapShot: NR coregulators. Cell **143**: 172–172.e1. doi:10.1016/j.cell.2010.09.032.
33. **Mootha VK, et al.** 2004. Erralpha and Gabpa/b specify PGC-1alpha-dependent oxidative phosphorylation gene expression that is altered in diabetic muscle. Proc. Natl. Acad. Sci. U. S. A. **101**:6570–6575.
34. **Mootha VK, et al.** 2003. PGC-1alpha-responsive genes involved in oxidative phosphorylation are coordinately downregulated in human diabetes. Nat. Genet. **34**:267–273.
35. **Nuutila P, et al.** 1994. Different alterations in the insulin-stimulated glucose uptake in the athlete's heart and skeletal muscle. J. Clin. Investig. **93**:2267–2274.
36. **Oka S, et al.** 2011. PPARalpha-Sirt1 complex mediates cardiac hypertrophy and failure through suppression of the ERR transcriptional pathway. Cell Metab. **14**:598–611.
37. **Perissi V, Aggarwal A, Glass CK, Rose DW, Rosenfeld MG.** 2004. A corepressor/coactivator exchange complex required for transcriptional activation by nuclear receptors and other regulated transcription factors. Cell **116**:511–526.
38. **Petersen KF, Dufour S, Befroy D, Garcia R, Shulman GI.** 2004. Impaired mitochondrial activity in the insulin-resistant offspring of patients with type 2 diabetes. N. Engl. J. Med. **350**:664–671.
39. **Potthoff MJ, et al.** 2007. Histone deacetylase degradation and MEF2 activation promote the formation of slow-twitch myofibers. J. Clin. Investig. **117**:2459–2467.
40. **Reilly SM, et al.** 2010. Nuclear receptor corepressor SMRT regulates mitochondrial oxidative metabolism and mediates aging-related metabolic deterioration. Cell Metab. **12**:643–653.
41. **Santos GM, Fairall L, Schwabe JW.** 2011. Negative regulation by nuclear receptors: a plethora of mechanisms. Trends Endocrinol. Metab. **22**:87–93.
42. **Schuler MJ, Buhler S, Pette D.** 1999. Effects of contractile activity and hypothyroidism on nuclear hormone receptor mRNA isoforms in rat skeletal muscle. Eur. J. Biochem. **264**:982–988.
43. **Schuler MJ, Pette D.** 1998. Quantification of thyroid hormone receptor isoforms, 9-cis retinoic acid receptor gamma, and nuclear receptor corepressor by reverse-transcriptase PCR in maturing and adult skeletal muscles of rat. Eur. J. Biochem. **257**:607–614.
44. **Shearer BG, et al.** 2008. Identification and characterization of a selective peroxisome proliferator-activated receptor beta/delta (NR1C2) antagonist. Mol. Endocrinol. **22**:523–529.
45. **Summermatter S, et al.** 2012. Remodeling of calcium handling in skeletal muscle through PGC-1alpha: impact on force, fatigability, and fiber type. Am. J. Physiol. Cell Physiol. **302**:C88–C99.
46. **Sutanto MM, et al.** 2010. The silencing mediator of retinoid and thyroid hormone receptors (SMRT) regulates adipose tissue accumulation and adipocyte insulin sensitivity in vivo. J. Biol. Chem. **285**:18485–18495.
47. **Suzuki H, et al.** 2009. The transcriptional network that controls growth arrest and differentiation in a human myeloid leukemia cell line. Nat. Genet. **41**:553–562.
48. **Webb P, et al.** 2000. The nuclear receptor corepressor (N-CoR) contains three isoleucine motifs (I/LXXII) that serve as receptor interaction domains (IDs). Mol. Endocrinol. **14**:1976–1985.
49. **Willy PJ, et al.** 2004. Regulation of PPARgamma coactivator 1alpha (PGC-1alpha) signaling by an estrogen-related receptor alpha (ERRalpha) ligand. Proc. Natl. Acad. Sci. U. S. A. **101**:8912–8917.
50. **Wisloff U, et al.** 2005. Cardiovascular risk factors emerge after artificial selection for low aerobic capacity. Science **307**:418–420.
51. **Yamamoto H, et al.** 2011. NCoR1 is a conserved physiological modulator of muscle mass and oxidative function. Cell **147**:827–839.
52. **Yoon HG, et al.** 2003. Purification and functional characterization of the human N-CoR complex: the roles of HDAC3, TBL1 and TBLR1. EMBO J. **22**:1336–1346.

Appendix 3

OPEN ACCESS Freely available online

PLoS one

PGC-1α Determines Light Damage Susceptibility of the Murine Retina

Anna Egger[1], Marijana Samardzija[2], Vithiyanjali Sothilingam[3], Naoyuki Tanimoto[3], Christina Lange[2], Silvia Salatino[1], Lei Fang[4], Marina Garcia-Garrido[3], Susanne Beck[3], Michal J. Okoniewski[5], Albert Neutzner[4], Mathias W. Seeliger[3], Christian Grimm[2], Christoph Handschin[1]*

1 Biozentrum, Division of Pharmacology/Neurobiology, Biozentrum, University of Basel, Basel, Switzerland, **2** Laboratory for Retinal Cell Biology, Department of Ophthalmology, University of Zurich, Schlieren, Switzerland, **3** Division of Experimental Ophthalmology, Institute for Ophthalmic Research, University of Tübingen, Tübingen, Germany, **4** Department of Biomedicine, University Eye Clinic, University Hospital Basel, Basel, Switzerland, **5** Functional Genomics Center UNI ETH Zurich, Zurich, Switzerland

Abstract

The peroxisome proliferator-activated receptor γ coactivator 1 (PGC-1) proteins are key regulators of cellular bioenergetics and are accordingly expressed in tissues with a high energetic demand. For example, PGC-1α and PGC-1β control organ function of brown adipose tissue, heart, brain, liver and skeletal muscle. Surprisingly, despite their prominent role in the control of mitochondrial biogenesis and oxidative metabolism, expression and function of the PGC-1 coactivators in the retina, an organ with one of the highest energy demands per tissue weight, are completely unknown. Moreover, the molecular mechanisms that coordinate energy production with repair processes in the damaged retina remain enigmatic. In the present study, we thus investigated the expression and function of the PGC-1 coactivators in the healthy and the damaged retina. We show that PGC-1α and PGC-1β are found at high levels in different structures of the mouse retina, most prominently in the photoreceptors. Furthermore, PGC-1α knockout mice suffer from a striking deterioration in retinal morphology and function upon detrimental light exposure. Gene expression studies revealed dysregulation of all major pathways involved in retinal damage and apoptosis, repair and renewal in the PGC-1α knockouts. The light-induced increase in apoptosis *in vivo* in the absence of PGC-1α was substantiated *in vitro*, where overexpression of PGC-1α evoked strong anti-apoptotic effects. Finally, we found that retinal levels of PGC-1 expression are reduced in different mouse models for *retinitis pigmentosa*. We demonstrate that PGC-1α is a central coordinator of energy production and, importantly, all of the major processes involved in retinal damage and subsequent repair. Together with the observed dysregulation of PGC-1α and PGC-1β in *retinitis pigmentosa* mouse models, these findings thus imply that PGC-1α might be an attractive target for therapeutic approaches aimed at retinal degeneration diseases.

Citation: Egger A, Samardzija M, Sothilingam V, Tanimoto N, Lange C, et al. (2012) PGC-1α Determines Light Damage Susceptibility of the Murine Retina. PLoS ONE 7(2): e31272. doi:10.1371/journal.pone.0031272

Editor: Richard Libby, University of Rochester, United States of America

Received November 11, 2011; **Accepted** January 4, 2012; **Published** February 13, 2012

Funding: This project was funded by the Swiss National Science Foundation, the SwissLife "Jubiläumsstiftung für Volksgesundheit und medizinische Forschung", the Roche Research Foundation, the United Mitochondrial Disease Foundation (UMDF) and the University of Basel. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing Interests: The authors have declared that no competing interests exist.

* E-mail: christoph.handschin@unibas.ch

Introduction

The vertebrate retina translates information of a photon of light into an electrical signal. It is a complex, 7-layered compartment at the innermost part of the eye and is made up of 8 distinct cell types, five of which are neuronal cells. Photons hit photosensitive cells (photoreceptors), which enable phototransduction by conformational change of the photopigment rhodopsin. Distinct photoreceptors allow for vision in light (cone cells) and dark (rod cells) conditions. Subsequent photopigment recovery occurs in the adjacent retinal pigment epithelium (RPE) [1]. The light information is transmitted from the photoreceptors via interneurons (amacrine cells, horizontal cells, bipolar cells) to ganglion cells, which then relay an electric action potential to the optical nerve and ultimately the visual cortex in the brain.

PGC-1 (peroxisome proliferator-activated receptor γ coactivator-1) designates a family of coactivators that comprises PGC-1α, PGC-1β and PGC-1-related coactivator (PRC). These coactivators dock to specific nuclear receptors and other transcription factors, thereby promoting the transcription of target genes, for example those implicated in mitochondrial biogenesis and oxidative phosphorylation (OXPHOS) [2–5]. PGC-1α and PGC-1β are mainly expressed in tissues with a high energetic demand, such as skeletal muscle, liver, pancreas, heart, kidney, brain and brown adipose tissue (BAT) [2–5]. The expression of PGC-1α and PGC-1β is further regulated in these organs by developmental stimuli and physiological stressors like cold, fasting and exercise. PGC-1α and PGC-1β accordingly regulate tissue-specific functions such as adaptive thermogenesis in brown adipose tissue, gluconeogenesis in the liver or endurance exercise adaptation in skeletal muscle [2–5].

Interestingly, despite the well-established link between PGC-1 expression and cellular energetics, the expression and function of the PGC-1 coactivators in the retina, one of the most energy

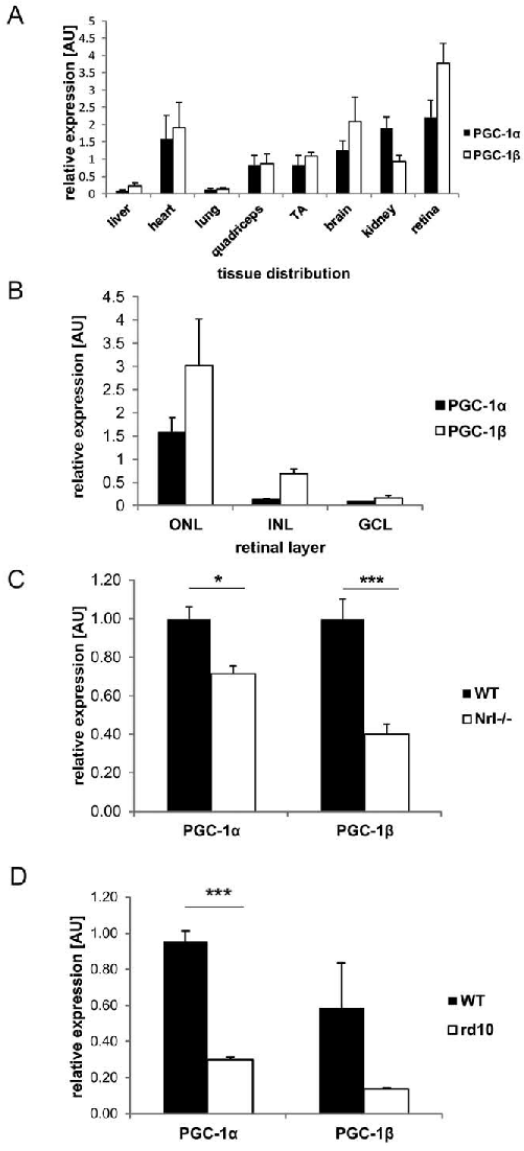


Figure 1. PGC-1α and PGC-1β are strongly expressed in photoreceptors in the retinal outer nuclear layer. Relative mRNA expression levels of PGC-1α and PGC-1β in different mouse tissues evaluated by semi- quantitative real time PCR. β- actin was used as housekeeping gene (HKG). P value was calculated using two tailed Student's T test: * = $p<0.05$; *** = $p<0.001$. +SEM **A.** Relative mRNA expression in different tissues from 3 month old wild type male C57BL/6j mice (TA = tibialis anterior). n = 4 **B.** Relative mRNA expression levels of PGC-1α and PGC-1β in different retinal layers from 3 month old wild type male C57BL/6 mice: ONL = outer nuclear layer, INL = inner nuclear layer, GCL = ganglion cell layer. n = 4 **C.** +**D.** Relative mRNA expression levels of PGC-1α and PGC-1β in two retinal disease models: Retina was isolated from *Nrl*−/−, rd10 and age matched control C57BL/6j WT mice. n = 3. **C.** Expression of PGC-1α and PGC-1β in *Nrl*−/− mice at 2 months of age. Data were normalized against C57BL/6j WT mice. **D.** Expression in rd10 mice at 5–6 months of age.
doi:10.1371/journal.pone.0031272.g001

demanding vertebrate organs [6], has not been studied so far, even though a number of cues warrant the analysis of this coactivator family in this organ. The high demand on ATP for physiological retinal function is assured by both glycolysis and oxidative phosphorylation: Neurotransmission relies on glycolysis, sodium transport employs glycolysis and oxidative phosphorylation, whereas phototransduction is mainly fuelled by oxidative phosphorylation to meet energetic demands [7]. In phototransduction, rods and cones are comparable in their energy expenditure in dark conditions, while cones have greater energy demands in the light [8]. Thus, most of the metabolic processes that are important for retinal function are most probably strongly regulated by the PGC-1 coactivators in analogy to other tissues, in particular substrate uptake, fatty acid oxidation and oxidative phosphorylation.

Inversely, pathological conditions of the retina such as age-related macular degeneration and diabetic retinopathy are associated with excess generation of reactive oxygen species (ROS), inflammation and endoplasmic reticulum stress, which result in apoptosis and tissue degeneration [9,10]. PGC-1α affects these processes by promoting ROS detoxification [11], ameliorating endoplasmic reticulum stress [12] and modulating tissue as well as systemic inflammation [13,14]. Finally, photoreceptor outer segment renewal involves activation of PPARγ [15], a functional interaction partner of the PGC-1 coactivators in various tissues [16].

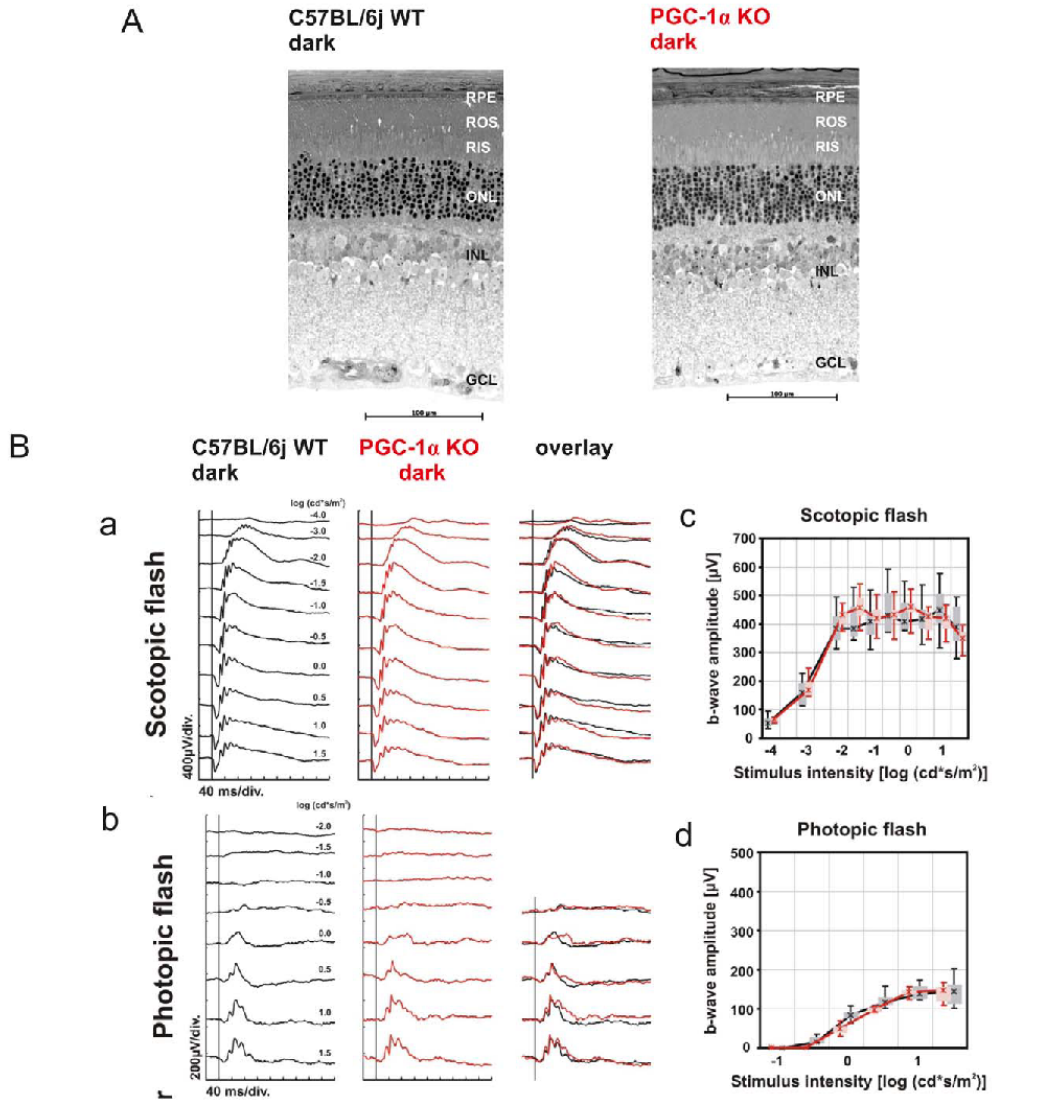
Thus, to investigate the potential role of PGC-1α in the retina, we assessed the expression patterns of PGC-1α and PGC-1β in retinal development, adult retina and mouse models for degenerative retinal diseases. Furthermore, we investigated the consequences of genetic ablation of PGC-1α in the retina in terms of changes in gene expression, morphology and function in the basal, dark condition and upon light-induced retinal stress and damage. Finally, we validated some of our findings in retinal cells *ex vivo* with acutely altered levels of PGC-1α. Our findings indicate a crucial role for PGC-1α in tissue protection, in part by inhibition of apoptosis.

Results

PGC-1α and PGC-1β are highly expressed in murine retina

Real-time qPCR-based expression analysis revealed highest PGC-1α and PGC-1β levels in tissues with a great energetic demand. Accordingly, PGC-1α and PGC-1β were found at elevated levels in skeletal and cardiac muscle, brain and kidney. Both coactivators were more moderately expressed in liver and lung. Strikingly, the highest expression levels of all tissues studied were observed in the retina for PGC-1α and PGC-1β (Fig. 1A).

To assess the differential expression pattern of PGC-1α and PGC-1β in distinct retinal compartments, we used laser capture microdissection to separate the different retinal cell layers and subsequently quantified expression via qPCR. PGC-1α and PGC-1β showed their highest expression in the outer nuclear layer (ONL) that harbors rod and cone photoreceptors (Fig. 1B). Although markedly lower compared to ONL, expression of both coactivators was also found in the inner nuclear layer (INL), which contains bipolar, horizontal, amacrine and Müller glia cells, and the ganglion cell layer (GCL), where ganglion and amacrine cells are located (Fig. 1B).



C

KO dark vs WT dark down	
pathway name	pValue
Oxidative phosphorylation	5.02E-16
Ubiquinone metabolism	4.8E-09
TCA	2.06E-05
Transcription_Transcription regulation of aminoacid metabolism	0.000839
Transcription_Role of VDR in regulation of genes involved in osteoporosis	0.002058
Aminoacyl-tRNA biosynthesis in cytoplasm	0.002839
Aminoacyl-tRNA biosynthesis in cytoplasm/ Rodent version	0.003044
Leucine, isoleucine and valine metabolism	0.005915
Cytoskeleton remodeling_Neurofilaments	0.006587
Unsaturated fatty acid biosynthesis	0.011499

KO dark vs WT dark up	
pathway name	pValue
Cytoskeleton remodeling_TGF, WNT and cytoskeletal remodeling	3.09E-07
Transcription_Androgen Receptor nuclear signaling	9.85E-07
Immune response_MIF - the neuroendocrine-macrophage connector	1.25E-06
Development_GM-CSF signaling	2.11E-05
Development_WNT signaling pathway. Part 2	3.6E-05
Development_Ligend-independent activation of ESR1 and ESR2	4.54E-05
Cytoskeleton remodeling_Cytoskeleton remodeling	4.58E-05
PGE2 pathways in cancer	5.03E-05
Development_TGF-beta-dependent induction of EMT via SMADs	5.2E-05
Development_Endothelin-1/EDNRA transactivation of EGFR	6.58E-05

Figure 2. Regular morphology and function in PGC-1α KO mouse retina. Dark adapted, 13 week old PGC-1α KO and C57BL/6j WT were subjected to morphological analysis and electroretinogram (ERG). **A.** Representative retinal morphologic sections from the lower central retina of 13 week old PGC-1α KO and C57BL/6j WT mice: RPE = retinal pigment epithelium, ROS = rod outer segments, RIS = rod inner segments, ONL = outer nuclear layer, INL = inner nuclear layer, GCL = ganglion cell layer. **B.** [**A**] Scotopic (dark-adapted) and [**B**] photopic (light-adapted) single-flash ERGs with increasing light intensities recorded from PGC-1α KO and C57BL/6j WT control mice. The vertical line crossing each trace shows the time point of the light flash. +SEM of n = 4. [**C**] Scotopic and [**D**] photopic b-wave amplitudes from PGC-1α KO and C57BL/6 control mice as a function of the logarithm of the flash intensity. Boxes indicate the 25% and 75% quartile range, whiskers indicate the 5% and 95% quantiles. **C.** Microarray analysis of PGC-1α KO and C57BL/6j WT control mice. Shown are the top 10 pathways up-/downregulated in their mRNA expression of PGC-1α KO and C57BL/6j WT mice, dark exposed. n = 3; * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$. Statistical significance was calculated using ANOVA test, Benjamin Hochberg corrected. Thresholds for changes in gene expression set at 1.2 (20% above control) or 0.80 (20% below control).
doi:10.1371/journal.pone.0031272.g002

To distinguish between rod and cone-specific expression of the PGC-1 coactivators within the ONL, we determined PGC-1α and PGC-1β levels in neural retinal leucine zipper (*Nrl*) knockout animals, which lack rod photoreceptors but retain cone morphology and function [17] and the rd10 mouse model where rod degeneration is followed by cone degeneration until all photoreceptors are lost by day 60 [18]. The significant reduction of PGC-1α and PGC-1β transcript levels in *Nrl* knockout mice (Fig. 1C) and the rd10 animal model (Fig. 1D) implies that both coactivators are expressed in rods and cones.

Genetic ablation of Pgc-1α does not compromise morphology and function in the dark adapted, unstressed mouse retina

PGC-1α global knockout animals [19] were used to study the function of PGC-1α in the retina. First, we analyzed basal morphology of 13 week- old PGC-1α knockout (KO) and C57 BL/6 WT control mice. Both KO and WT animals revealed an intact, regularly shaped retinal structure (Fig. 2A). To assess basal retinal function, the mice were kept in the dark and then subjected to an electroretinogram (ERG), which tracks the electrical responses of retinal cells to a light stimulus. The resulting a and b and waves reflect the function of the outer and inner layer, respectively. ERG responses were recorded in scotopic (darkness) and photopic (illuminated) conditions to analyze rod and cone function, respectively. In all of these experimental contexts, the PGC-1α KO mice showed regular a and b waves undistinguishable from the WT control animals (Fig. 2B). Finally, in addition to the morphological and functional studies, we investigated whether genetic ablation of the PGC-1α gene in the retina affects global gene expression patterns with gene expression microarrays. Gene ontological analysis revealed primarily a reduction in metabolic pathways in the KO animals, foremost oxidative phosphorylation and citrate cycle (Fig. 2C), similar to results in loss-of-function experiments with PGC-1α in other tissues. Inversely, cytoskeleton remodeling, transforming growth factor beta (*Tgfβ*) and *Wnt* signaling were the most prominent pathways expressed at a higher level in KO compared to WT mice (Fig. 2C).

Absence of PGC-1α increases light-damage susceptibility

While in most organs, the effects of genetic ablation on morphology and function are relatively mild, stress conditions like exercise, cold or fasting greatly exacerbate the phenotype of PGC-1α knockout mouse models [13,19]. Thus, triggered by our observation of a small effect of PGC-1α knockout on retinal morphology and function in the dark, we sought to study the repercussions of PGC-1α ablation after employing a physiological, retina-specific stressor. For this purpose, PGC-1α KO and WT control mice were exposed to 15.000 lux of white light for 2 hrs, which corresponds to direct sun light exposure. After a 24 hrs recovery period, morphologies of PGC-1α deficient and WT control retinae were analyzed. In stark contrast to the dark-adapted mice (Fig. 2), PGC-1α KO and WT control mice responded markedly differently to a strong light insult. First, most of the rod inner (RIS) and outer (ROS) segments were completely destroyed and most nuclei were pyknotic in the ONL of PGC-1α KO mice (Fig. 3A). Thus, most photoreceptors have lost their light absorption (ROS) and energy-production (RIS) sites. In contrast, the corresponding regions in WT exhibited only occasional pyknotic nuclei (Fig. 3A) and only one out of six wild type animals showed measurable retinal damage, which, however, was still less intense than in PGC-1α KOs (data not shown). In either of the genotypes, the INL and GCL layers were unharmed consistent with previous data showing that light damage causes selective death of photoreceptor cells [20] (Fig. 3A).

Given the nature and extent of retinal degeneration observed in the KO animals, it seemed highly likely that these morphological changes would translate into loss of retinal function. Light-exposed PGC-1α KO mice displayed a blunted scotopic ERG compared to the WT mice with a significant reduction in b wave amplitude 10 days post insult (Fig. 3B). This timeframe was chosen as retina clears injured photoreceptors and returns back to a basal state within 10 days [21]. The photopic response in KO animals remained similar to WT mice, which is consistent with the notion that cones remain unharmed in an acute light- damage paradigm, at least at early timepoints after light damage [22].

Subsequent to the ERG analysis, animals were subjected to spectral domain optical coherence tomography (SD-OCT), which allow for highly sensitive cross sectional analysis of retinal layers, respectively, and thus reveal retinal detachment as well as delicate structural impairments. Cross sectional retinal analysis showed that the KO mice that had a reduced scotopic ERG response also displayed a thinning of the ONL compared to WT controls. (Fig. 3C- abc). This thinning was also observed in the morphological analysis in Fig. 3A. Interestingly however, the morphology indicates an altered arrangement of the nuclei in the ONL to underlie the thinning. Thus, since the total number of nuclei in KO and WT animals is indistinguishable, an increase in relative density is found in the thinner ONL of KO mice (KO: 43 nuclei/mm^2, WT: 31 nuclei/mm^2).

Light-exposed PGC-1α KO animals exhibit reduced expression of phototransduction genes, increased expression of apoptotic and inflammatory markers

To elucidate the deficits in gene expression in the light-stressed PGC-1α KO animals, we compared gene expression microarrays of WT and KO mice in the control, dark-adapted state and after light exposure, respectively. Interestingly, the influence of the environmental conditioning from dark to light exposure was similar in the two genotypes in regard to the number of regulated genes. However, most of the 2914 genes regulated in the KO animals between dark and light were different from the 2524 genes in the WT animals or were modulated at a significantly higher level. Thus, light- exposed PGC-1α KO and WT mice differed in the expression of 1774 genes, 897 of which were elevated and 877 of which were decreased in KO compared to the WT mice, more

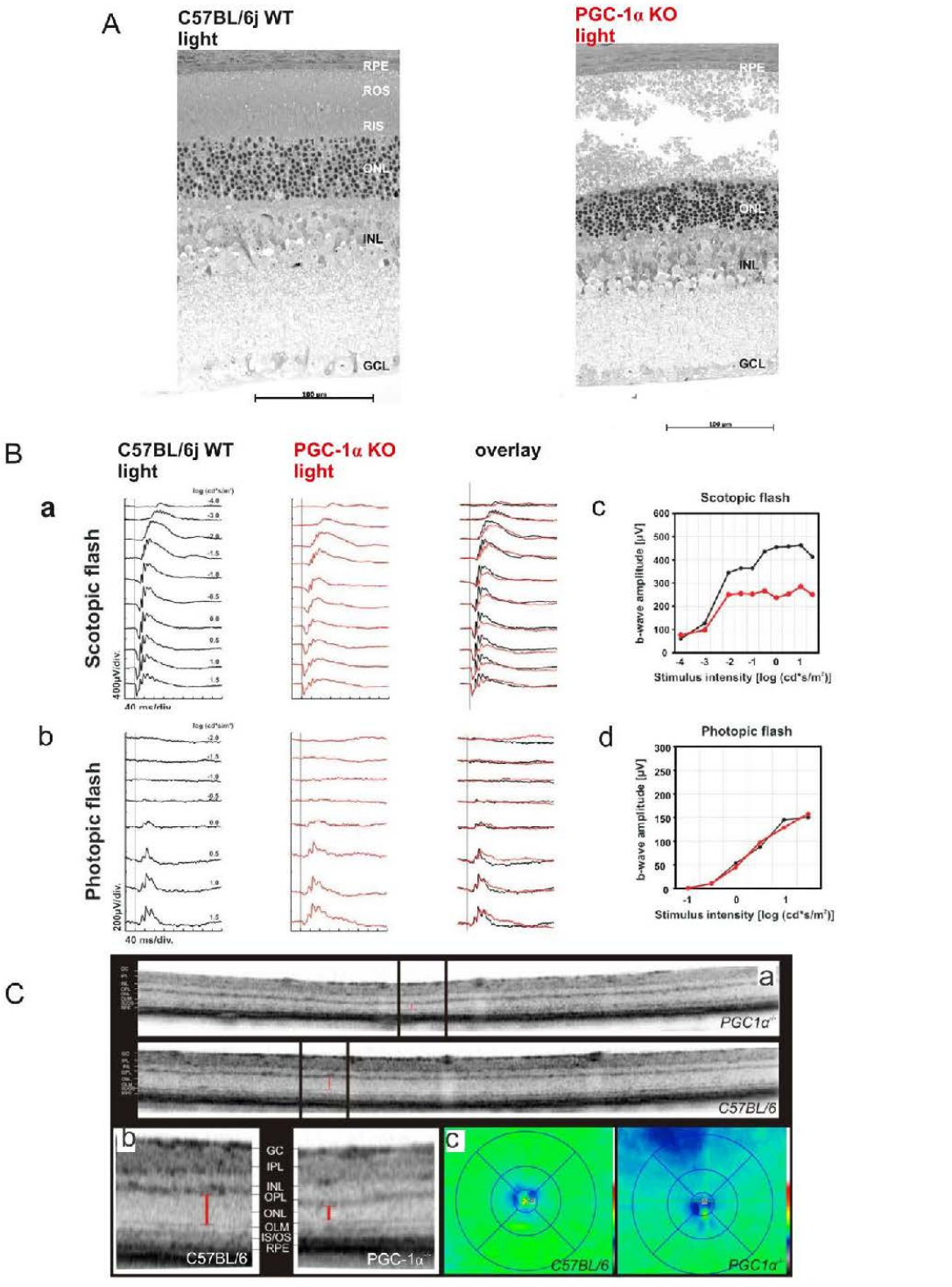

A
C57BL/6j WT
light
PGC-1α KO
light
RPE
ROS
RIS
ONL
INL
GCL
B
C57BL/6j WT
light
PGC-1α KO
light
overlay
a
Scotopic flash
c
Scotopic flash
b-wave amplitude [µV]
Stimulus intensity [log (cd*s/m²)]
b
Photopic flash
d
Photopic flash
b-wave amplitude [µV]
Stimulus intensity [log (cd*s/m²)]
C
a
PGC1α⁻/⁻
C57BL/6
b
c
GC
IPL
INL
OPL
ONL
OLM
IS/OS
RPE
C57BL/6
PGC-1α⁻/⁻
C57BL/6
PGC1α⁻/⁻

Figure 3. Aberrant morphology and compromised function of PGC-1α KO mouse retina under light stress conditions. PGC-1α *KO* and C57BL/6j WT, aged 13 weeks, were exposed to 15.000 lux of white light for 2 hours and after 24 hours and 10 days of recovery subjected to morphological and functional analysis, respectively. **A.** Representative retinal morphologic sections from the lower central retina of PGC-1α KO and C57BL/6j WT mice. **B.** [**A**] Scotopic (dark-adapted) and [**B**] photopic (light-adapted) single-flash ERGs with increasing light intensities recorded from PGC-1α KO and C57BL/6j WT control mice. The vertical line crossing each trace shows the time point of the light flash. [**C**] Scotopic and [**D**] photopic b-wave amplitudes from PGC-1α KO and C57BL/6 control mice as a function of the logarithm of the flash intensity. (**c**) Optical coherence tomography (SD-OCT): [**A**] comparison of two representative OCT horizontal slices of PGC-1α KO and C57BL/6j WT control mice. [**B**] Magnification of these SD-OCT horizontal slices: left side: C57BL/6, right side: PGC-1α KO; GC = ganglion cell layer, IPL = inner plexiform layer, INL = inner nuclear layer, OPL = outer plexiform layer, ONL = outer nuclear layer, OLM = outer limiting membrane, IS/OS = inner segment/outer segment, RPE = retinal pigment epithelium [**C**] Topography of retinal thickness of C57BL/6 (side) and PGC-1α KO (right) mice calculated from 61 equidistant SD-OCT slices (green = dense; blue = less dense).
doi:10.1371/journal.pone.0031272.g003

than three times the number of genes as observed in the corresponding dark accommodated mouse comparison (Fig. 4A, Fig. S1E). The top regulated genes (Table S1A) and the most affected pathways (Table S1B) that were downregulated in KO mice compared to WT animals still included metabolic processes as observed in the dark condition. Strikingly however, various key processes in retinal function, damage and repair, in particular visual perception, photoreceptor function, DNA damage and repair were now lower in the KO animals compared to controls (Fig. 4B, 4C and Fig. S1A–C). Phototransduction pathway analysis showed that PGC-1α KO mice displayed diminished expression of phototransduction genes in the illuminated state (for example KO light vs WT light: *Nrl, Nr2e3, Arr3, Scn4a, Cnga1*) (Fig. 4B, 4C, Table S1B). Similarly, the expression of many anti-apoptotic genes was reduced in the animals with an ablated PGC-1α gene, including *Fas* (Fig. 4B, 4C and Table S1). Inversely, many inflammatory and pro-apoptotic genes were expressed at a higher level in light-exposed KO mice compared to controls, such as *Bcl2* and *A2M*, respectively (Fig. 4B, 4C). Interestingly, genes involved in extrinsic or intrinsic apoptosis were upregulated in the PGC-1α KO vs WT animals; however, no specific apoptotic pathway was preferentially affected (Table S2A). Finally, only the WT animals exhibited an increase in oxidative phosphorylation gene expression pattern comparing dark adapted to light exposed mice (Table S1B: WTli vs WTda up; KOli vs KOda up).

PGC-1α expression alleviates apoptosis in retinal cells

Light exposure of the retina leads to damage and induction of apoptotic pathways. Based on our observation in light exposed animals, we tested whether PGC-1α function is directly involved in the regulation of early events in photoreceptor apoptosis and we therefore quantified cell death in PGC-1α KO and WT control animals following light exposure. In this experimental context, the generation of free nucleosomes was detected in both genotypes, but at significantly higher levels in the PGC-1α KO than in the WT control animals (Fig. 5A). Thus, the deterioration of the KO retina upon light exposure involves an increase in apoptotic events.

To circumvent the potentially confounding effects of a global, life-long absence of PGC-1α in the KO mice that may be compensated *in vivo*, we investigated the effect of acutely modulated PGC-1α in ARPE-19 cells, a retinal pigment epithelial cell line. In these cells, we studied cytochrome *c* release as an early event in intrinsic, mitochondria-related apoptosis. ARPE-19 cells were transduced with adenovirus encoding GFP as control or PGC-1α expression cassettes, respectively. Importantly, despite the induction of cytochrome *c* expression by PGC-1α that has been reported in other experimental contexts (Fig. 5B), no change in cytochrome *c* release was observed in ARPE-19 cells with ectopic elevation of PGC-1α (Fig. 5C). As expected, treatment of the ARPE-19 cells with actinomycin D or camptothecin caused elevated cytochrome *c* release and thus apoptosis in GFP-expressing control cells. However, when the cells were additionally expressing PGC-1α, this resulted in a significant reduction in cytochrome *c* release that was statistically undistinguishable from cytochrome *c* release in vehicle- treated control cells. (Fig. 5B).

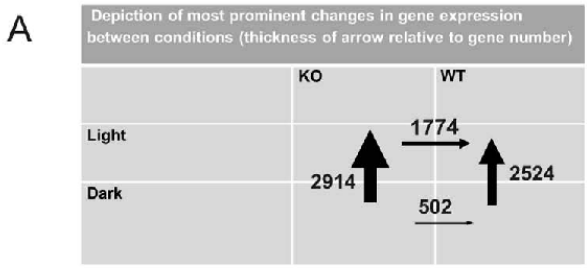
PGC-1α and PGC-1β are reduced in rd10 and VPP models of degenerative retinal disease

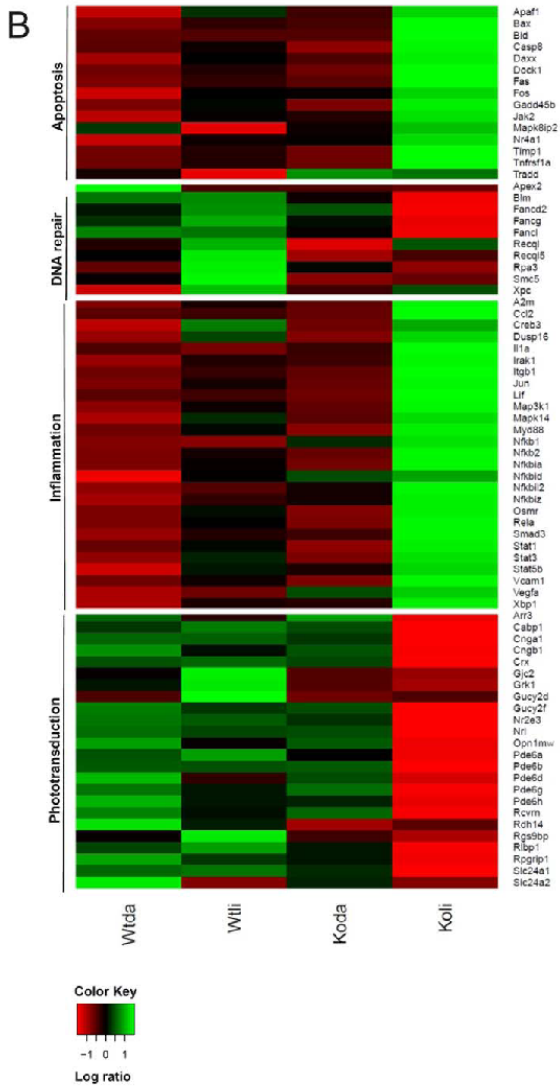
A wide array of retinal disease conditions is associated with increased apoptosis of photoreceptors, including *retinitis pigmentosa*, a multi- genetic hereditary retinal disease that is caused by mutations in genes implicated in photoreceptor and RPE function. *Retinitis pigmentosa* (RP) causes progressive vision loss by the degeneration of rods and cones [23]. Animal models to study RP include the rd10, the VPP and the *Nrl* KO mouse [18,24–25]. Since apoptotic and other pathways involved in retinal damage in *retinitis pigmentosa* were affected by the PGC-1α gene knockout in light- exposed mice, we investigated whether the expression levels of PGC-1α were altered in these mouse models with age-dependent disease progression of retinal degeneration. In the WT animals, PGC-1α augmented with increasing age similar to *Rho (Rhodopsin)*, a marker of rod differentiation. In contrast, after an initial rise, PGC-1α expression reverted back to initial levels in the rd10 mouse (Fig. 6A, 6B). A similar decrease was observed for PGC-1β expression in the rd10 animals, where PGC-1β levels remained consistently lower compared to those of WT animals, even after the initial spike of PGC-1β expression around the age of 15 days (Fig. 6C). Any remaining amount of both coactivators was probably due to their presence in the GCL and INL which remained intact in all of the used models (Fig. 1B).

The VPP mouse suffers from a degeneration of rods and cones caused by three rhodopsin mutations (V20G, P23H, and P27L [VPP]) [24]. Both the rd10 and the VPP mouse models thus recapitulate the damage in rods and cones that is observed in human *retinitis pigmentosa*. Interleukin-1β (*Il-1β*), served as a control gene that is upregulated in the VPP mouse due to the inflammation in this model [26] (Fig. 6D), similar to what was shown for the inflammatory marker monocyte chemotactic protein 1(*Mcp-1*) [27]. Both PGC-1α and PGC-1β were expressed at a lower level in the VPP mouse compared to WT animals. Strikingly, convergence of both PGC-1α and PGC-1β expression in VPP and WT animals coincided with the relief of the inflammatory response in the VPP model implied by the strong decrease in *Il-1β* expression between day 42 and day 52. Thus, the apoptotic events seen in the VPP model that precede the inflammation directly correlated with reduced levels of PGC-1α and PGC-1β (Fig. 6D, 6E, 6F). Interestingly, PGC-1β was likewise reduced in the light- exposed PGC-1α knockout mice, which also exhibited retinal damage and dysfunction, as shown in the present manuscript (Fig. S1E).

Discussion

The PGC-1 coactivators are key regulators of mitochondrial biogenesis and oxidative metabolism. Hence, our present


A
Depiction of most prominent changes in gene expression between conditions (thickness of arrow relative to gene number)
KO
WT
Light
Dark
1774
2914
2524
502


B
Apoptosis
DNA repair
Inflammation
Phototransduction
Wtda
Wtli
Koda
Koli
Color Key
−1 0 1
Log ratio

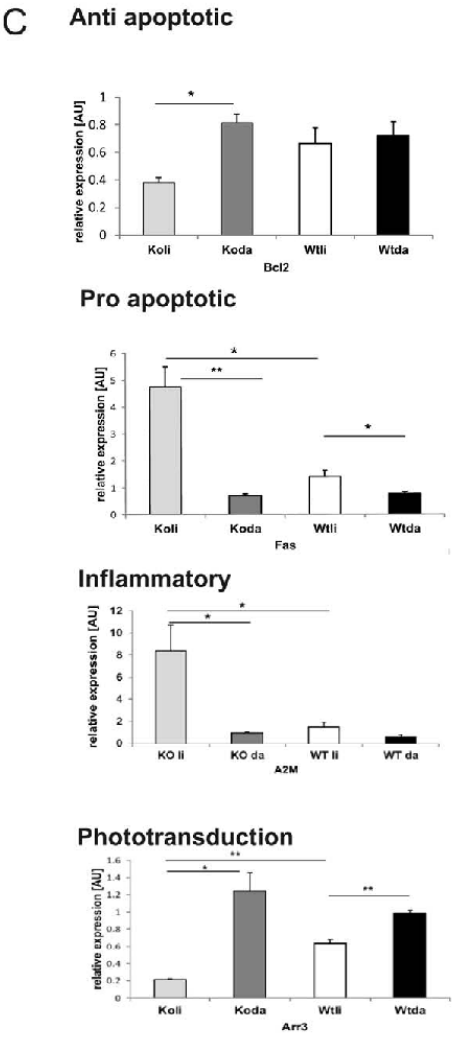

C
Anti apoptotic
relative expression [AU]
Koli
Koda
Wtli
Wtda
Bcl2
Pro apoptotic
Fas
Inflammatory
KO li
KO da
WT li
WT da
A2M
Phototransduction
Arr3

Figure 4. Microarray: Ablation of PGC-1α affects gene expression changes relative to apoptosis, inflammation, protein folding and phototransduction. Microarray analysis of light exposed PGC-1α KO and C57BL/6j WT control mice. +SEM for n = 3 **A.** Number of genes changed significantly between genotypes and conditions: thickness of arrow relates to number of genes changed. **B.** Heat map: expression patterns of a representative selection of genes involved in apoptosis, inflammation, DNA repair and phototransduction. Green = upregulated; red = downregulated; black = unchanged. Data were analyzed with the PARTEK and Bioconductor software suites using Benjamin-Hochberg-corrected ANOVA tests in Panels A and B. Threshold of significant gene expression changes were set at 1.2 fold up- and 0.87 fold down-regulation, respectively. **C.** Semi-quantitative real time PCR validation examples of pro apoptotic, anti apoptotic, inflammatory and phototransduction genes changed significantly in the microarray. 18SrRNA was used as HKG. P value was calculated using two tailed Student's T test: * = $p<0.05$; ** = $p<0.01$. +SEM of n = 3; Koli = PGC-1α knockout, light condition, Koda = PGC-1α KO, dark condition, Wtli = C57BL/6 WT, light conditions, Wtda = C57 BL/6 WT, dark condition.
doi:10.1371/journal.pone.0031272.g004

observation of high expression of both PGC-1α and PGC-1β in the retina, a tissue with one of the highest energy demands, is not surprising. Moreover, the intraretinal expression pattern showing high levels of both coactivators in rod and cone photoreceptors reflects the high energetic requirement of phototransduction that to a large extent is met by oxidative phosphorylation [7]. Interestingly however, dark-adapted mice with an ablation of the PGC-1α gene in the retina display normal retinal morphology and function, even despite a notable reduction in metabolic gene expression in this experimental context. Lack of a prominent phenotype in PGC-1α loss-of-function settings in basal, unstressed conditions has also been observed in other organs in tissue-specific knockout mice [28,13,29] and in global knockout animals [19,30]. Importantly, a more striking phenotypic dysregulation in these animal models is often induced by external stimuli such as fasting, cold exposure or exercise. Similarly, the absence of a functional PGC-1α gene has much more dramatic consequences in light-exposed compared to dark-adapted animals. Light stress and the resulting cellular damage in the retina results in a cascade of events aimed at destruction of damaged cells through apoptosis, removal of cellular debris, repair of DNA damage and restoration of phototransduction through increased expression of cone- and rod-specific genes (Fig. 6G). These processes require a considerable amount of energy, witnessed in the increased expression of genes implicated in oxidative phosphorylation in the WT light versus WT dark up (Table S1B). For example, components of the NADH dehydrogenase (*Ndufa*) and the *cytochrome c* oxidase (*Cox*) complexes are upregulated in this context. The PGC-1α KO animals, however, likely fail to meet the increased energy need (Table S1: KO li vs WT li down) and thus restoration of tissue homeostasis and function is hampered (Fig. 6G). In addition, apoptosis and the inflammatory response are enhanced in light-damaged retinae of PGC-1α knockout animals, while DNA damage repair and photoreceptor recovery are impaired in these mice (Table S2). Interestingly, a recent study analogously implied that PGC-1α dysregulation may be implicated in irradiation-caused DNA damage in the liver [31]. As a consequence, light-exposed retinae of PGC-1α knockout animals show morphological aberrations, specifically a destruction of the RIS and ROS. Interestingly, the thinning of the ONL in the KO animals seems to be primarily caused by a reduction in cell size and/or spacial cellular arrangement since the number of nuclei in the thinner ONL in the KO is not reduced, at least not at the time of the morphological analysis. This interesting observation will be the focus of future studies. As a result of the impairment in morphology and gene expression, light perception is significantly impaired in these mice, as suggested in particular by the reduced the rod-dependent, scotopic b-wave amplitude. In contrast, photopic light perception is unaltered: this difference between the drastic morphological phenotype and the apparently milder functional outcome could be due to the experimental setup. First, since mice were able to freely move during the light exposure, damage was not homogenous throughout the whole retina. Second, morphological and functional analysis were not performed at the same time after light exposure and were assessed in separate animals. Nevertheless, taken together, our data indicate an important role for PGC-1α in the protection against light-induced retinal damage, in providing energy for light perception and in repair and renewal of the photoreceptors of the retina.

The interpretation of our data obtained in global PGC-1α knockout animals could potentially be hampered by several drawbacks of this model. First, total ablation of the PGC-1α gene in many cases results in different, sometimes diametrically opposed phenotypic changes compared to those in tissue-specific knockouts [32]. For example, hepatic gluconeogenesis is constitutively elevated in one global PGC-1α knockout mouse model [19], whereas the fasting-dependent induction of gluconeogenesis is significantly blunted in liver-specific knockouts [29]. Similarly, fiber-type distribution in global knockout is unchanged compared to control mice, while AMP-dependent protein kinase (AMPK) is activated in this context [19,33]. In contrast, a fiber-type switch towards glycolytic muscle fibers has been observed in muscle-specific PGC-1α knockouts that also exhibit normal AMPK activation [28,13]. In both cases, the tissue-specific knockout models exhibit the phenotype that is expected based on *in vivo* gain-of-function experiments [34,35]. Second, in a global PGC-1α knockout mouse, PGC-1α is obviously not only absent in retinal cells, but also other structures involved in light perception and vision. For example, PGC-1α knockout animals suffer from marked neurodegenerative events in the central nervous system [19]. Neuronal deficits, in particular in the visual cortex where PGC-1α expression has been reported [36], but also the optic nerve, would obviously also affect vision in these mice. Finally, it is difficult to interpret etiology and causality underlying the pathological phenotype in the global PGC-1α knockout animals with a life-long absence of a functional PGC-1α gene. To circumvent all of these potentially confounding problems of global, non-inducible PGC-1α gene ablation and to test whether apoptosis is directly affected by PGC-1α and not altered as a consequence of an energy crisis, we tested whether acute modulation of PGC-1α levels in an isolated system of RPE cells recapitulates our observations made in light-exposed PGC-1α knockout mice exhibiting increased levels of apoptosis *in vivo*. Our experiments in the ARPE-19 cell line clearly show a strong anti-apoptotic effect of ectopically expressed PGC-1α against two different pharmacological triggers of apoptosis, actinomycin D and camptothecin. PGC-1α expression has been associated with decreased apoptosis in different cellular contexts, including skeletal muscle [37], aortic endothelial cells [38] and decreased retinal capillary cell apoptosis after statin treatment [10]. In contrast, PGC-1α regulates intestinal epithelial cell fate in part by promoting apoptosis in the intestinal epithelial surface [39]. We now show for the first time that loss-of-function of PGC-1α results in increased apoptosis in the retina *in vivo* and that increased levels of PGC-1α exert an anti-apoptotic effect in retinal pigment epithelial cells.

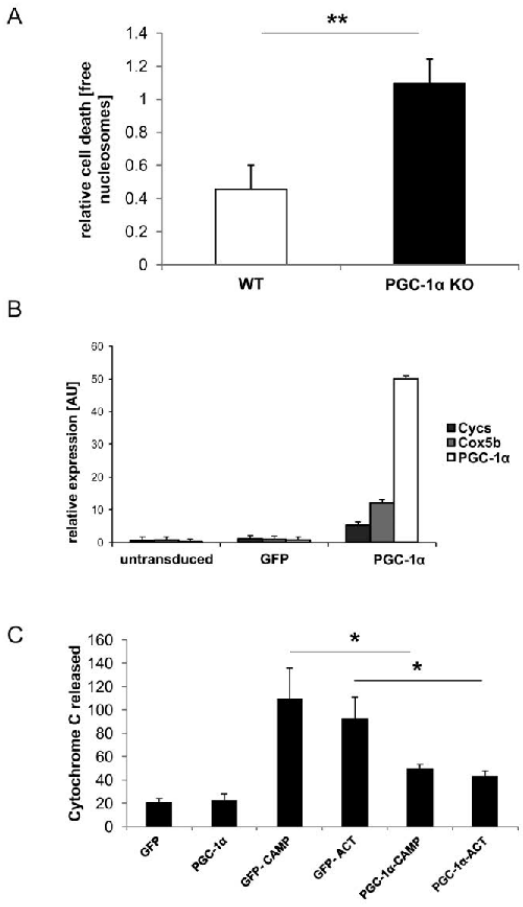


Figure 5. PGC-1α alleviates retinal apoptotic response upon light stress in vivo and in vitro. Relative quantification of cell death. P value was calculated using two tailed Student's T test: * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$ **A.** PGC-1α KO and C57BL/6j WT mice were exposed to 15'000 lux of white light for 2 hrs. After 24 hrs, retina was extracted and analyzed for free nucleosomes as a measure of cell death via ELISA. +SEM of n = 11 (WT) and n = 15 (PGC-1α KO). **B.** Retinal pigment epithelial cells ARPE-19 were transduced with adenovirus carrying GFP or mouse PGC-1α. Semi-quantitative real time PCR of PGC-1α and its targets *Cox5b* (cytochrome c oxidase subunit Vb) and *Cycs* (cytochrome *c*) in PGC-1α versus GFP transduced ARPE-19 cells. 18SrRNA was used as HKG. +SEM of n = 6 **C.** ARPE-19 cells were transduced with adenovirus carrying GFP or PGC-1α, treated with apoptosis inducers actinomycin D or camptothecin and then analyzed for cell death via immunostaining of released cytochrome c. ACT = Actinomycin D, CAMP = Campthothecin. 18SrRNA was used as HKG. +SEM of n = 6.

doi:10.1371/journal.pone.0031272.g005

Our tissue culture data thus not only confirm the direct involvement of loss-of-function of PGC-1α on apoptosis in the ONL *in vivo*, but together with the observation of reduced PGC-1α and PGC-1β expression in mouse models for *retinitis pigmentosa* also indicate that elevation of PGC-1α might be a valid approach to prevent retinal damage and also promote recovery in retinal diseases through increase in ATP generation. Future studies should thus aim at elucidating the therapeutic efficacy of normalization of PGC-1α levels with viral vectors or pharmacological tools on retinal degeneration in *retinitis pigmentosa* and other retinal pathologies.

Methods

Animals

All of the animal experiments have been approved by the institution and authorities of the cantons Basel-Stadt and Zurich in Switzerland as well as the institutional and federal authorities in Germany (no. 114/2010) and performed according to the statement of "The Association for Research in Vision and Ophthalmology" for the use of animals in research. Wildtype C57 BL/6 NRj mice were purchased from Janvier (Le Genest St. Isle, France). C57 BL/6 PGC-1α global KO mice were a kind gift of Bruce M. Spiegelman, Dana- Farber Cancer Institute and Department of Cell Biology, Harvard Medical School (Boston, MA). Rd10 (= retinal degeneration 10) mice were purchased from Jackson Lab (Bar Harbor, Maine). VPP (= mutant opsin transgene) mice were obtained from Muna Naash, University of Oklahoma (Oklahoma City, Oklahoma). *Nrl* (Neural retinal leucine zipper) KO mice were purchased from the University of Michigan (Ann Arbor, Michigan). Mice were maintained in a 12 h:12 h light-dark cycle (60 lux).

Light exposure

Six to thirteen- week old male C57 BL/6 PGC-1α global KO mice and wildtype C57 BL/6 NRj mice were dark adapted overnight (16 hrs). Pupils were dilated with 1% cyclogyl (Alcon, Cham, Switzerland) and 5% phenylephrine (Ciba Vision, Niederwangen, Switzerland) for 45 min. before light exposure. Mice were exposed to 15'000 lux of light for two hours. Following overnight recovery in darkness, mice were analyzed at time points indicated in the results section. Mice that were dark adapted and not exposed to light served as controls.

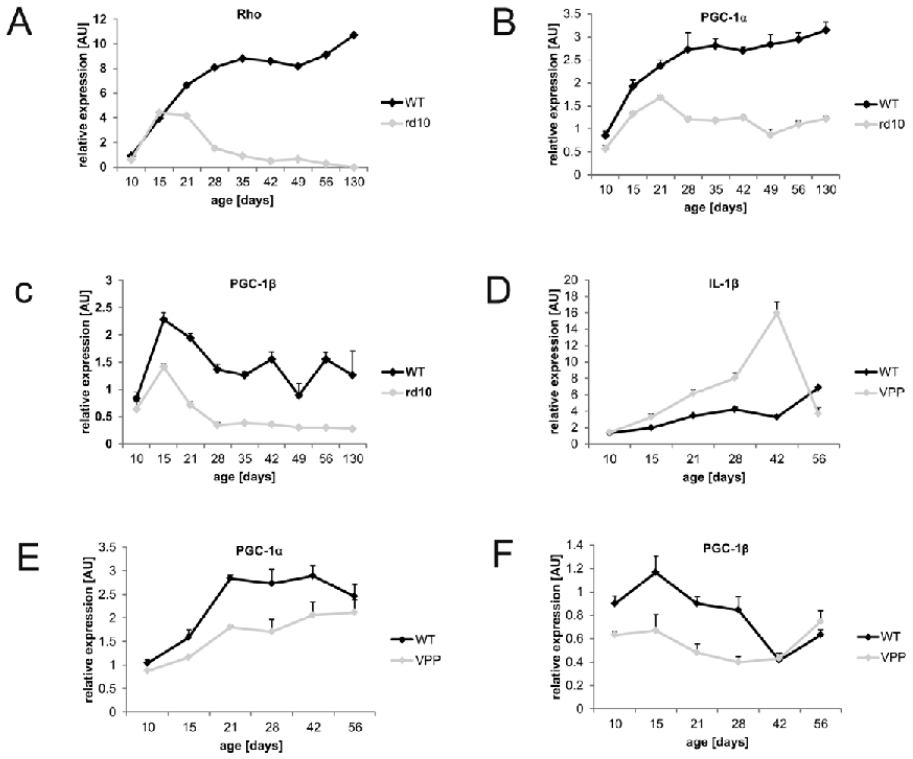
Morphology and cell death detection ELISA

Following a 24 hours recovery period after light exposure, eyes were fixed in 2.5% glutaraldehyde in 0.1 M cacodylate buffer (pH 7.3) at 4°C overnight as a preparation for morphological analysis. The superior and inferior retinae were prepared, washed in cacodylate buffer, incubated in osmium tetroxide for 1 hour, dehydrated graded in alcohol series. Epon 812 sections (0.5 µM) from the lower central retina were counter-stained with methylene blue.

Cell death was quantified 24 hours after light exposure in isolated retinas using the ELISA-based cell death detection kit (Roche Diagnostics, Mannheim, Germany). Free nucleosomes, a hallmark of apoptosis in the retina and other tissues, were quantified using anti-histone antibodies in the KO and WT retinae. This method allows quantification of internucleosomal degradation of genomic DNA during apoptosis in proliferating and non-proliferating cells while excluding necrotic DNA.

Laser capture microdissection

Mice were sacrificed, eyes enucleated, embedded in tissue freezing medium (Leica Microsystems Nussloch, Germany) and frozen in 2-methylbutane bath cooled by liquid nitrogen. Retinal sections (20 µM) were fixed (5 min. acetone), air-dried (5 min.) and dehydrated (30 sec 100% ethanol, 5 min. xylene). Microdissection was performed using an Arcturus XT Lasercapture device (Molecular devices, Silicon Valley, CA).


A
Rho
relative expression [AU]
age [days]
WT
rd10
B
PGC-1α
C
PGC-1β
D
IL-1β
VPP
E
PGC-1α
F
PGC-1β

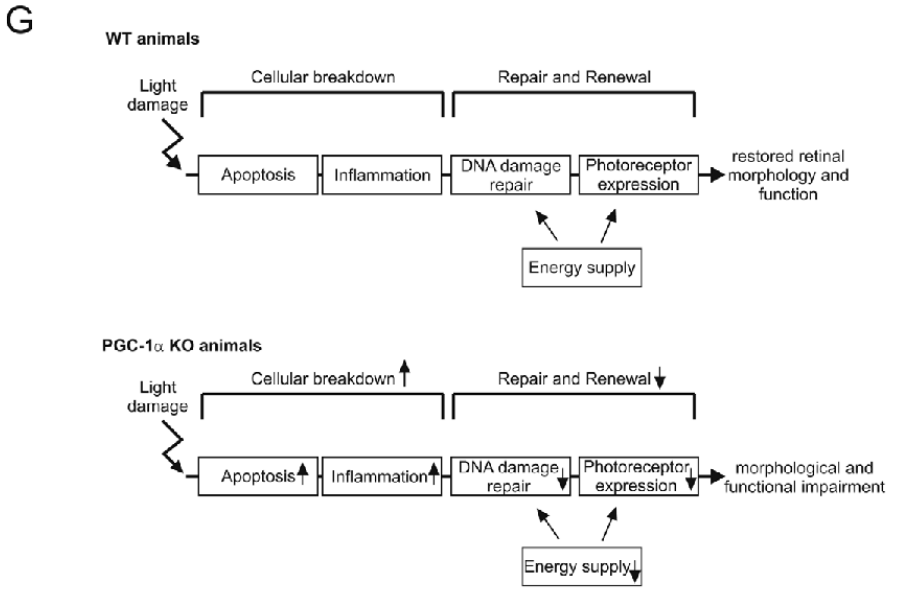

G
WT animals
Light damage
Cellular breakdown
Repair and Renewal
Apoptosis
Inflammation
DNA damage repair
Photoreceptor expression
restored retinal morphology and function
Energy supply
PGC-1α KO animals
morphological and functional impairment

Figure 6. PGC-1α and PGC-1β are reduced in rd10 and VPP retinal degeneration models. Retina from rd10 and VPP mice was isolated at different PNDs (postnatal days) and checked for expression of genes via semi- quantitative real time PCR. β-actin was used as HKG. +SEM of n = 3 per time point. **A–C** rd10 mouse model: Rhodopsin (*Rho*) was used as control gene diminished in this condition. **D–F** VPP mouse model: Interleukin 1-β (*Il-1β*) was used as control gene marking the duration of the inflammatory response phase. **G** Loss of the PGC-1α gene in mice affects all the steps of the major cascade of events in light-induced retinal damage, repair and regeneration and consequently results in abnormal retinal morphology and function.
doi:10.1371/journal.pone.0031272.g006

Gene Expression Studies

RNA was isolated by pooling 2 retinas from 1 animal using the miRNeasy Mini Kit from Qiagen. 1 µg of RNA was employed for cDNA synthesis using Super Script Reverse Transcriptase II (Invitrogen) and random hexamer primers (Promega, Madison, USA). Real-time PCR analysis (Power SYBR Green Master Mix, Applied Biosystems, Carlsbad, California, USA) was performed either by using the LightCycler 480 Sybr Green I Master Kit and a LightCycler 480 instrument (Roche, Mannheim, Germany) or by Power SYBR Green Master Mix using the StepOnePlus sequence detector (Applied Biosystems, Carlsbad, California, USA). Relative expression levels were calculated with the ΔΔCt method and normalized to the expression of β-actin or 18S ribosomal RNA (18SrRNA) with primers indicated in Table S3.

Cell culture and adenoviral infection

ARPE-19 cells were seeded at 2×10E5 cells per well of a 12 well plate in 50% F-12/50% DMEM medium supplemented with 1% Penicillin/Streptomycin. 24 h afterward, cells were infected with adenoviral constructs encoding GFP and PGC-1α, respectively.

Apoptosis induction and cell death detection

48 h after adenoviral infection, cells were treated with Z-VAD-FMK (Peptanova, Sandhausen, Germany) for 60 min to prevent cell detachment. Apoptosis inducers Camptothecin or Actinomycin D (final concentration: 20 µM) were added. 24 h later, cells were fixed with 4% paraformaldehyde in PBS, washed with PBS and stained with anti-cytochrome *c* (BD biosciences). Released cytochrome *c* was quantified using fluorescence microscopy.

ERG measurements

Electroretinograms (ERGs) were recorded according to previously described procedures [40]. Mice that were light exposed as described above were analyzed 10 days after light damage. Non-exposed mice served as controls. All mice were dark- adapted overnight prior to ERG analysis. Mice were then anaesthetized with ketamine (66.7 mg/kg body weight) and xylazine (11.7 mg/kg body weight). Pupils were dilated and single flash ERG responses were obtained under dark-adapted (scotopic) and light-adapted (photopic) conditions. Light adaptation was accomplished with a background illumination of 30 candela (cd) per square meter starting 10 min before recording. Single white-flash stimulus intensity ranged from −4 to 1.5 log cd·s/m2 under scotopic and from −2 to 1.5 log cd·s/m2 under photopic conditions.

Spectral domain optical coherence tomography (SD-OCT)

Eyes were subjected to SD-OCT using the commercially available Spectralis™ HRA+OCT device from Heidelberg Engineering featuring a broadband superluminescent diode at λ = 870 nm as low coherent light source [41].

Gene expression profiling

RNA was isolated as described above (Gene expression studies) and 4.25 ug of cDNA from the mouse retinae was hybridized to the Affymetrix Gene Chip Mouse 1.0 ST at the in-house microarray facility service. Microarray CEL output files were analyzed at the Functional Genomics Center Zurich, Winterthurerstrasse 190, CH-8057 Zurich. Statistical significance was calculated using the ANOVA test, Benjamin Hochberg corrected with the PARTEK and Bioconductor software suites. Thresholds for changes in gene expression set at 1.2 (0.2× upregulated) or 0.87 (0.2× downregulated). $p \leq 0.05 = *$, $p \leq 0.01 = **$, $p \leq 0.001 = ***$.

Supporting Information

Figure S1 Gene expression analysis of WT and PGC-1α KO mice in light and dark. (**a**) Microarray analysis of light and dark exposed PGC-1α KO and C57BL/6j WT control mice; Gene expression changes in the four comparisons: 18sRNA was used as HKG. +SEM for n = 3; * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$. Statistical significance was calculated using two tailed Student's T test (**b**) phototransduction genes, (**c**) pro and anti-apoptotic genes, (**d**) inflammatory genes and (**e**) PGC-1α and PGC-1β gene expression.
(PDF)

Table S1 Microarray analysis of light and dark exposed PGC-1α KO and C57BL/6j WT control mice: top genes and pathways. (**a**) Top 10 genes most highly up-/downregulated in their mRNA expression of PGC-1α KO and C57BL/6j WT mice, dark vs light exposed. (**b**) Top 10 pathways up-/downregulated in their mRNA expression of PGC-1α KO and C57BL/6j WT mice, dark vs light exposed. n = 3; * = $p<0.05$; ** = $p<0.01$; *** = $p<0.001$. Statistical significance was calculated using ANOVA test, Benjamin Hochberg corrected. Thresholds for changes in gene expression set at 1.2 (0.2× upregulated) or 0.87 (0.2× downregulated).
(PDF)

Table S2 Microarray analysis of light and dark exposed PGC-1α KO and C57BL/6j WT control mice. Significant changes in genes involved in apoptosis, DNA repair, inflammation and phototransduction. green = upregulated; red = downregulated (**a**) pro and anti-apoptotic, DNA repair genes (**b**) phototransduction and protein folding genes (**c**) ECM breakdown and inflammatory genes.
(PDF)

Table S3 List of primers for semi-quantitative real time PCR analysis.
(PDF)

Acknowledgments

We thank our colleagues in the laboratory for discussions and comments on the manuscript.

Author Contributions

Conceived and designed the experiments: AE MWS CG CH. Performed the experiments: AE MS CL VS SB MG-G NT LF AN CG. Analyzed the data: AE MWS MO SS CG CH. Wrote the paper: AE CH.

References

1. Lamb TD, Pugh EN (2006) Phototransduction, dark adaptation, and rhodopsin regeneration the proctor lecture. Invest Ophthalmol Vis Sci 47: 5137–5152.
2. Lin J, Handschin C, Spiegelman BM (2005) Metabolic control through the PGC-1 family of transcription coactivators. Cell Metab 1: 361–370.
3. Handschin C, Spiegelman BM (2006) Peroxisome proliferator-activated receptor gamma coactivator 1 coactivators, energy homeostasis, and metabolism. Endocr Rev 27: 728–735.
4. Finck BN, Kelly DP (2006) PGC-1 coactivators: inducible regulators of energy metabolism in health and disease. J Clin Invest 116: 615–622.
5. Handschin C (2009) The biology of PGC-1alpha and its therapeutic potential. Trends Pharmacol Sci 30: 322–329.
6. Nathans J (1987) Molecular biology of visual pigments. Ann Rev Neurosci 10: 2897–2908.
7. Ames A, Li YY, Heher EC, Kimble CR (1992) Energy metabolism of rabbit retina as related to function: high cost of Na+ transport. J Neurosci 12: 840–853.
8. Okawa H, Sampath AP, Laughlin SB, Fain GF (2008) ATP Consumption by Mammalian Rod Photoreceptors in Darkness and Light. Curr Biol 18: 1917–1921.
9. Herzlich AA, Ding X, Shen D, Ross RJ, Tuo J, et al. (2009) Peroxisome proliferator-activated receptor and age-related macular degeneration. Open Biol J 2: 141–148.
10. Zheng Z, Chen H, Wang H, Ke B, Zheng B, et al. (2010) Improvements of retinal vascular injury in diabetic rats by statins is associated with the inhibition of mitochondrial reactive oxygen species pathway mediated by peroxisome-proliferator- activated receptor gamma coactivator 1 alpha. Diabetes 59: 2315–2325.
11. St-Pierre J, Drori S, Uldry M, Silvaggi JM, Rhee J, et al. (2006) Suppression of reactive oxygen species and neurodegeneration by the PGC-1 transcriptional coactivators. Cell 127: 397–408.
12. Wu J, Ruas JL, Estall JL, Rasbach KA, Choi JH, et al. (2011) The unfolded protein response mediates adaptation to exercise in skeletal muscle through a PGC-1α/ATF6α complex. Cell Metab 13: 160–169.
13. Handschin C, Chin S, Li P, Liu F, Maratos-Flier E, et al. (2007) Skeletal muscle fiber-type switching, exercise intolerance, and myopathy in PGC-1alpha muscle-specific knock-out animals. J Biol Chem 282: 30014–30021.
14. Handschin C, Spiegelman BM (2008) The role of exercise and PGC-1alpha in inflammation and chronic disease. Nature 454: 463–469.
15. Ershov AV, Bazan NG (2000) Photoreceptor phagocytosis selectively activates PPARgamma expression in retinal pigment epithelial cells. J Neurosci Res 60: 328–337.
16. Puigserver P, Wu Z, Park CW, Graves R, Wright M, et al. (1998) A cold-inducible coactivator of nuclear receptors linked to adaptive thermogenesis. Cell 92: 829–839.
17. Mears AJ, Kondo M, Swain PK, Takada Y, Bush RA, et al. (2001) Nrl is required for rod photoreceptor development. Nat Genet 29: 447–452.
18. Chang B, Hawes NL, Pardue MT, German AM, Hurd RE, et al. (2007) Two mouse retinal degenerations caused by missense mutations in the beta-subunit of rod cGMP phosphodiesterase gene. Vision Res 47: 624–633.
19. Lin J, Wu P-H, Tarr PT, Lindenberg KS, St-Pierre J, et al. (2004) Defects in adaptive energy metabolism with CNS-linked hyperactivity in PGC-1alpha null mice. Cell 119: 121–135.
20. Grimm C, Wenzel A, Hafezi F, Yu S, Redmond TM, et al. (2000) Protection of Rpe65-deficient mice identifies rhodopsin as a mediator of light-induced retinal degeneration. Nat Genet 25: 63–66.
21. Wenzel A, Grimm C, Samardzija M, Remé CE (2005) Molecular mechanisms of light-induced photoreceptor apoptosis and neuroprotection for retinal degeneration. Prog Retin Eye Res 24: 275–306.
22. Organisciak DT, Vaughan DK (2010) Retinal Light Damage: Mechanisms and Protection. Prog Retin Eye Res 29: 113–134.
23. Sahel J, Bonnel S, Mrejen S, Paques M (2010) Retinitis pigmentosa and other dystrophies. Dev Ophthalmol 47: 160–167.
24. Naash MI, Hollyfield JG, al-Ubaidi MR, Baehr W (1993) Simulation of human autosomal dominant retinitis pigmentosa in transgenic mice expressing a mutated murine opsin gene. Proc Natl Acad Sci U S A 90: 5499–5503.
25. Daniele LL, Lillo C, Lyubarsky AL, Nikonov SS, Philp N, et al. (2005) Cone-like morphological, molecular, and electrophysiological features of the photoreceptors of the Nrl knockout mouse. Invest Ophthalmol Vis Sci 46: 2156–2167.
26. Samardzija M, Wenzel A, Thiersch M, Frigg R, Remé C, et al. (2006) Caspase-1 ablation protects photoreceptors in a model of autosomal dominant retinitis pigmentosa. Invest Ophthalmol Vis Sci 47: 5181–5190.
27. Joly S, Lange C, Thiersch M, Samardzija M, Grimm C (2008) Leukemia inhibitory factor extends the lifespan of injured photoreceptors in vivo. J Neurosci 28: 13765–13774.
28. Handschin C, Choi CS, Chin S, Kim S, Kawamori D, et al. (2007) Abnormal glucose homeostasis in skeletal muscle – specific PGC-1α knockout mice reveals skeletal muscle – pancreatic β cell crosstalk. J Clin Invest 117: 3463–3474.
29. Handschin C, Lin J, Rhee J, Peyer A-K, Chin S, et al. (2005) Nutritional regulation of hepatic heme biosynthesis and porphyria through PGC-1alpha. Cell 122: 505–515.
30. Leone TC, Lehman JJ, Finck BN, Schaeffer PJ, Wende AR, et al. (2005) PGC-1alpha deficiency causes multi-system energy metabolic derangements: muscle dysfunction, abnormal weight control and hepatic steatosis. PLoS Biol 3: e101.
31. Kim H, Jee HJ, Yun J (2011) DNA damage induces down-regulation of PEPCK and G6P gene expression through degradation of PGC-1a. Acta Biochim Biophys Sin (Shanghai) 43: 589–594.
32. Handschin C, Spiegelman BM (2011) PGC-1 coactivators and the regulation of skeletal muscle fiber-type determination. Cell Metab 13: 351.
33. Arany Z, He H, Lin J, Hoyer K, Handschin C, et al. (2005) Transcriptional coactivator PGC-1 alpha controls the energy state and contractile function of cardiac muscle. Cell Metab 1: 259–271.
34. Lin J, Wu H, Tarr P, Zhang C, Wu Z, et al. (2002) Transcriptional co-activator PGC-1α drives the formation of slow-twitch muscle fibres. Nature 418: 797–801.
35. Yoon JC, Puigserver P, Chen G, Donovan J, Wu Z, et al. (2001) Control of hepatic gluconeogenesis through the transcriptional coactivator PGC-1. Nature 413: 131–138.
36. Yu L, Yang SJ (2010) AMP-activated protein kinase mediates activity-dependent regulation of peroxisome proliferator-activated receptor gamma coactivator-1alpha and nuclear respiratory factor 1 expression in rat visual cortical neurons. Neuroscience 169: 23–38.
37. Adhihetty PJ, Uguccioni G, Leick L, Hidalgo J, Pilegaard H, et al. (2009) The role of PGC-1alpha on mitochondrial function and apoptotic susceptibility in muscle. Am J Physiol Cell Physiol 297: C217–C225.
38. Won JC, Park J-Y, Kim YM, Koh EH, Seol S, et al. (2010) Peroxisome proliferator-activated receptor-gamma coactivator 1-alpha overexpression prevents endothelial apoptosis by increasing ATP/ADP translocase activity. Arterioscler Thromb Vasc Biol 30: 290–297.
39. D'Errico I, Salvatore L, Murzilli S, Lo Sasso G, Latorre D, et al. (2011) Peroxisome proliferator-activated receptor-gamma coactivator 1-alpha (PGC1alpha) is a metabolic regulator of intestinal epithelial cell fate. Proc Natl Acad Sci U S A 108: 6603–6608.
40. Tanimoto N, Muehlfriedel R, Fischer M, Fahl E, Humphries P, et al. (2009) Vision tests in the mouse: Functional phenotyping with electroretinography. Front Biosci 14: 2730–2737.
41. Fischer MD, Huber G, Beck SC, Tanimoto N, Muehlfriedel R, et al. (2009) Noninvasive, in vivo assessment of mouse retinal structure using optical coherence tomography. PloS ONE 4: e7507.